



CENTRO DE INVESTIGACIONES BIOLÓGICAS DE MADRID

Desarrollo y validación experimental de un método para la obtención de estructuras de complejos ligando/receptor basado en Resonancia Magnética Nuclear

Javier Rodríguez Salarichs

27/06/2017

Dirigida por:

Dr. José Fernando Díaz Pereira

Dra. Ángeles Canales Mayordomo

Memoria para optar por el grado de Doctor en Biofísica en la Universidad
Autónoma de Madrid.

Dedicado a la memoria de mis padres



Mirtha Salarichs García

Justo Rodríguez Giral



Agradecimientos

En un trabajo de tantos años se termina debiendo gratitud a muchas de las personas que se conocen, todas ellas, en su día a día te han podido aportar desde un valioso grano de arena a todo un saco lleno. Las palabras se pueden hacer cortas o incluso insulsas al intentar expresar lo agradecido que estoy de todos ustedes que, por curiosidad, me están leyendo. Todos ustedes, que esperan un espacio en este apartado, quiero que sepan que ya tiene ese espacio conquistado en mis recuerdos. Además, me gustaría decirles que siempre podrán contar con mi voluntad de ayudarlos, como yo en algún instante pudo contar con la de ustedes.

Primero que todo, me gustaría agasajar con este trabajo y con mi eterno agradecimiento a mis padres. A ellos va dedicado este trabajo y mi eterno amor como su hijo. La parte de la vida que me tocó vivir a sus lados fue especial y curiosa. Ellos se encargaron de crear las bases para que hoy pudiera estar escribiendo estas palabras. ¡Gracias por todo!

Mi hija, mi niñita, mi nena, mi princesa, gracias por ser parte de mi vida y espero que en algún momento puedas leer este texto. Si tu amor llega a tanto y lees estas líneas, pues que sepas que sin importar donde esté o lo que haga, siempre te llevo conmigo. Te quiero ante todo mi retoño de mí.

A mi flor tenerifeña le debo mi felicidad en todo momento. Sin importar lo dura o negra que puede ser la vida, a su lado y a través de su prisma, veo “La vida en rosa”. Por último, me gustaría agradecerle por haber sido mi editora personal.

A mi hermanita del alma, a mi abu querida y a Luquita, son tan especiales en mi vida que mejor que agradecerles aquí con palabras será decirles en persona ;-).

Fer, sabes que eres importante para mí, me adoptaste en cierta manera y en todo momento me apoyaste como un padre. Así que te agradezco por todo el apoyo desinteresado que me diste y por poder contar contigo.

Ángeles, fuiste la primera han acogerme. Gracias por dejar hacer y deshacer. Debido a esta libertad y a tu paciencia nacieron las primeras líneas del código de programa que ahora presento.

Mariano, otra persona importante, al hombre que más cafés le voy a deber en todo lo que me queda de vida ☺. Además, también fue mi editor personal, ¡j!, que dura tarea no?! ¡Muchas gracias por todo!

Gon, estuve compartiendo con él unos muy buenos años del doctorado. Es una persona muy difícil de olvidar. ¡¡¡Esto, claro está, se debe principalmente a lo que molestaba con sus constantes preguntas!!! ☺. Un especial abrazo para ti.

Yana, Adan y Erney, casi mis hermanos y mi hermana, gracias por todo. A la pandilla de Cuba, olé!!!

Dear Rama Krishna, I would like to thank you for all your concerns.

Al resto que no he nombrado, espero que sepan que están en mi corazón.

¡Gracias!

I am grateful to the UAB Research Foundation (UABRF), the National Institutes for Health (NIH) and the National Science Foundation (NSF) of USA for providing license of CORCEMA-ST.

Índice de contenidos

1	Introducción	9
1.1	Difracción de Rayos-X.....	10
1.2	Resonancia Magnética Nuclear	14
1.2.1	Efecto Nuclear Overhauser	19
1.2.2	Espectroscopia bidimensional	22
1.2.3	Efecto Nuclear Overhauser transferido.....	24
1.2.4	Diferencia de Transferencia de Saturación.....	33
1.3	Caracterización de la unión colvalente entre la tubulina y la dactilolida 43	
2	Hipótesis	53
3	Objetivos.....	53
4	Materiales y métodos.....	57
4.1	Productos químicos.....	57
4.2	Tampones	57
4.3	Proteínas.....	57
4.4	Preparación de tubulina en condiciones no polimerizantes.....	59
4.5	Preparación de muestras para experimentos de RMN.....	59
4.6	Experimentos de RMN	60
4.7	Experimentos de cristalografía y DRX.....	60
4.8	Cálculos de modelado molecular	61
4.9	Cálculos con el CORCEMA-ST	62
5	Resultados.....	65
5.1	Diseño de funciones de coste para la comparación de los STD-NMR experimentales y los calculados.....	65

5.1.1	Diseño de alternativas al factor R_e	65
5.1.2	Búsqueda de sistemas modelos para la evaluación de las funciones diseñadas	71
5.1.3	Estudio de la variación del patrón de STD-NMR de los sistemas modelos con los parámetros químico-físicos	73
5.1.4	Estudio de la variación de los factores R-NOE con la variación de los parámetros químico-físicos	80
5.1.5	Análisis cuantitativo del efecto de los parámetros químico-físicos en los factores R-NOE	86
5.1.6	Análisis de la variación de los factores R-NOE con el refinamiento estructural	94
5.1.7	Evaluación de los factores R-NOE en el reconocimiento de estructuras conocidas ligando-receptor	106
5.2	Diseño e implementación del STD-MaPa, software para la determinación estructural de complejos moleculares con datos de RMN ...	113
5.2.1	Antecedentes	113
5.2.2	Diseño de la metodología	114
5.2.3	Componentes del software STD-MaPa	116
5.2.4	Complejos moleculares de referencia.....	128
5.2.5	Evaluación de la calidad de los factores R-NOE en el ajuste combinado de parámetros y estructuras de los complejos	135
5.2.6	Resolución de la estructura de los complejos problemas mediante el programa STD-MaPa	142
5.2.7	Cálculo de múltiples modos de unión con el STD-MaPa	146
5.3	Método híbrido de RMN y DRX para la resolución estructural de complejos proteína-ligando	156
5.3.1	Resolución de la estructura del complejo dactilolida-tubulina.....	156

5.3.2	Cristalografía del complejo dactilolida-tubulina.....	161
5.3.3	Medición de TR-NOESY y STD-NMR del complejo dactilolida-tubulina	165
5.3.4	Resolución de la estructura del complejo con STD-NMR.....	169
5.3.5	Resolución de las estructuras de los múltiples modos de unión del complejo dactilolida- β -tubulina.....	175
6	Discusión	187
6.1	Nuevos factores R-NOE.....	187
6.2	STD-MaPa es capaz de generar estructuras válidas a pesar de que la aproximación CORCEMA no contemple los efectos de reasociación y de la relajación T1	194
6.3	STD-MaPa en la resolución de constantes químico-físicas	211
6.4	Descripción del mecanismo de unión covalente de la dactilolida y sus derivados a la β -tubulina	216
6.5	Resolución de la estructura de complejos moleculares con estructuras similares por STD-MaPa	221
7	Conclusiones	231
8	Conclusions	235
9	Anexos.....	241
9.1	Ecuaciones diferenciales de los factores R_t y R_e	241
9.2	Pruebas estadísticas de Levene, Barlett y Shapiro-Wilk.....	244
9.3	Publicaciones	247
10	Bibliografía	257

Índice de ilustraciones

Figura 1.1.1 Representación del fenómeno de difracción de rayos-X en un sólido cristalino (A). Corrimiento de fase que da lugar a las interferencias constructivas entre las ondas según la ley de Bragg (B).....	11
Figura 1.2.1 Representación de la situación de equilibrio de los espines nucleares con $I=1/2$ en ausencia de un campo magnético externo (izquierda) y cuando están bajo la influencia de un campo magnético H_0 (derecha). En la zona inferior de la figura se muestra los niveles de energía en ausencia y presencia de H_0 respectivamente.....	16
Figura 1.2.2 Modelo vectorial de Dirac y los diferentes acoplamientos entre los espines de las partículas atómicas.....	18
Figura 1.2.3 Diagrama de transiciones de dos espines nucleares aislados entre los cuatro estados de espines posibles, cuando se considera un sistema de núcleos con $I=1/2$, no acoplados escalarmente y en ausencia de otro tipo de interacción que no sea dipolar. w_i para $i=0,1,2$ es la probabilidad de transición entre estados.....	20
Figura 1.2.4 Espectroscopia de correlación. Consiste en un simple pulso de radiofrecuencia (p_1), seguido por un tiempo de evolución (t_1) y un posterior pulso de radiofrecuencia (p_2). Después del último pulso se inicia el período de medida (t_2).	22
Figura 1.2.5 Representación esquemática de los espectros bidimensionales homonucleares COSY (izquierda) y NOESY (derecha). En la zona superior de la imagen está representada la estructura de la molécula hipotética a partir de la cual se obtienen los dos espectros. En azul se muestra el acoplamiento escalar y en rojo las interacciones NOE.....	23
Figura 1.2.6 Esquema básico de la espectroscopia 2D de intercambio descrita en el trabajo de J. Jeener (Jeener, Meier et al. 1979).	25

Figura 1.2.7 Ilustración de la variación del efecto NOE en función del tiempo de correlación y la frecuencia del espectrómetro. 26

Figura 1.2.8 Izquierda, representación esquemática de un espectro NOESY. Los picos diagonales y los de cruce tienen diferentes signos. Derecha, representación esquemática de un espectro TR-NOESY. Los picos diagonales y los de cruce tienen los mismos signos. El tamaño relativo de los picos se escogió respetando la cercanía de los protones en el estado enlazado y libre. La superficie azul representa la proteína. 28

Figura 1.2.9 Secuencia de pulsos descrita por Bernd Meyer and Thomas Peters para el experimento STD-NMR (Meyer and Peters 2003). El pulso selectivo se compone de un pulso gaussiano de 50 ms y un tiempo de espera entre pulsos de $\delta=1$ ms. La intensidad del pulso gaussiano de saturación selectiva es $\gamma B_1=86$ Hz. n es el número de pulsos y típicamente es igual a 40. La frecuencia del pulso selectivo es modificada al finalizar cada secuencia de pulsos entre -1,5 ppm (magnetización de los protones metálicos de la proteína) y 100 ppm. d_1 es el tiempo que se espera entre las distintas acumulaciones para que las señales relajen. ϕ_1 y ϕ_2 son las fases de los pulsos. t_2 es el tiempo de adquisición. $T_{1\rho}$ es un pulso de bloque de espín ("*spin-lock*") de 30 ms y fase ϕ_2 que se utiliza para eliminar las señales anchas debidas a la proteína. En este procedimiento de STD-NMR, la substracción se realiza al finalizar las secuencias de pulsos..... 35

Figura 1.2.10 Representación esquemática de los diferentes pasos de un experimento STD-NMR. En la parte superior de la imagen se encuentran representadas las uniones entre la proteína y el ligando en un experimento estándar de ^1H -RMN (izquierda) y cuando se satura selectivamente los protones de la proteína (derecha). Además, se ilustran los procesos de relajación del sistema: intercambio químico (rojo), relajación intermolecular (naranja) y relajación intramolecular en el ligando (verde). La relajación intramolecular en la proteína no se representa. En la parte inferior de la figura se ilustran los espectros que se obtendrían a partir de los dos pasos. Además,

se muestra el espectro diferencia (STD-NMR). La saturación protónica se representa con tonalidades de azules, siendo representados los protones más saturados con un azul intenso y los protones menos saturados con un azul claro. En ausencia de saturación los protones se representan con un morado claro. 37

Figura 5.3.1. Estructura de la dactilolida. (A-B) Micrografías del efecto celular de la dactilolida sobre la red de microtúbulos y la morfología nuclear en el ciclo celular de la célula de carcinoma de pulmón A549. Micrografías de células A549 incubadas por 24h con DMSO (A) y 10 μ M de dactilolida (B). Los microtúbulos (verde) se muestran mediante la inmunotinción con el anticuerpo monoclonal de α -tubulina y el DNA (azul) se muestra mediante su tinción con Hoechst 33342. Insertados, el huso mitótico en las mismas preparaciones. La barra de escala representa 10 μ m. (C) Representación de un ensayo de competición entre el Flutax-2 y MSAs por la unión por los microtúbulos. Desplazamiento de flutax-2 (50 μ M) por zampanolida (negra), docetaxel (rojo), epotilona-A (verde) y dactilolida (azul). (D) Cinética de reacción de diferentes ligandos con los microtúbulos. Círculos, Flutax-2 remanente en el sitio de unión de Taxol cuando se incubaba con DMSO (negro), ZMP natural (azul claro), ZMP sintética (verde) y DAC (rojo). Cuadrados, DAC sin reaccionar presente en solución en presencia de tampón (negro) y DMSO (rojo). 46

Figura 5.1.1. Estructuras químicas de los ligandos de (A) kifunesina, (B) salacinol y (C) epotilona-A. Datos de STD-NMR experimentales y calculados a partir de la estructura cristalográfica de los complejos de (D) kifunesina-dGMII (PDB: 1PS3), (E) salacinol-dGMII (PDB: 1TQS) y (F) epotilona-A- α , β -heterodímero de tubulina (PDB: 4I50). Los datos de STD-NMR experimental están representados con líneas y círculos negros. Los datos de STD-NMR calculados están representados con líneas y cuadrados rojos. 72

Figura 5.1.2. Representación de la variación del STD-NMR promedio experimental y calculado del complejo kifunesina-dGMII (A y D), salacinol-dGMII (B y E) y epotilona A- α , β -heterodímero de tubulina (C y F). El promedio de cada complejo molecular se calculó a partir de las intensidades de los

protones del ligando saturados tomando como referencia el promedio experimental. El área sombreada y las barras de errores representan la desviación cuadrática media de los datos experimentales y teóricos respectivamente. N_H es el número de protones saturados por ligando. 76

Figura 5.1.3 Promedio normalizado de los factores R-NOE en función de los tiempos de correlación del ligando unido ($\tau_{c-bound}$ o τ_c^p) a distintos valores de tiempo de correlación del ligando libre (τ_{c-free} o τ_c^l). La normalización se realizó tomando como referencia los valores máximos (R_{max}) y ajustándolos a una función exponencial. El tiempo de correlación del ligando libre aparece señalado en la primera gráfica y mantiene el mismo código de colores para todas las demás. En la esquina superior derecha de cada gráfica se encuentran los factores R-NOE en función de los tiempos de correlación de los ligandos libres para los mejores ajustes del STD-NMR experimental promedio. 83

Figura 5.1.4 Factores R-NOE normalizados en función de la constante de equilibrio (K_a) a diferentes valores de velocidad de disociación del ligando (k_{off}) para tres ligandos diferentes (kifunesina, salacinol y epotilona). Los factores fueron normalizados tomando el valor máximo (R_{max}) para cada estudio como referencia. El comportamiento de los factores frente a las velocidades de asociación fue ajustado a una función exponencial y conserva el mismo código de colores para las distintas funciones de la figura. En la esquina superior derecha de cada gráfica se encuentran insertados los factores R-NOE en función de la velocidad de disociación cuando la constante de asociación presenta el mejor ajuste con el STD-NMR experimental promedio. 85

Figura 5.1.5. Diagrama de caja y bigote de la variación máxima normalizada de los factores R-NOE en función de los parámetros de las matrices K (ecuación 1.2.9) y R (ecuación 1.2.8). En la parte superior se encuentra una tabla con la media de los valores de los seis factores R-NOE y sus desviaciones estándar. 87

Figura 5.1.6. Exploración conformacional de la kifunesina, el salacinol y la ephotilone-A enlazados a sus proteínas dianas. (A-C) Representación de la estructuras cristalográficas de la kifunesina unida a dGMII (A), salacinol unido a dGMII (B) y la ephotilone-A unida al complejo T₂R (C). Los átomos que forman los planos diedros están resaltados en amarillo. (D-F) Representación del STD-NMR promedio normalizado en la exploración conformacional de la kifunesina (D), el salacinol (E) y la ephotilone-A (F). (G-I) Representación de la energía torsional descrita por el campo de fuerzas OPLS en la exploración conformacional de la kifunesina (G), el salacinol (H) y la ephotilone-A (I). Los puntos señalan las conformaciones encontradas por los métodos de DRX (rojo), las diferencias entre las intensidades experimentales y calculadas de STD-NMR (amarillo) y el campo de fuerzas OPLS2005 (verde). Además, se muestran los valores de RMSD entre las conformaciones..... 97

Figura 5.1.7. Representación de las superficies conformacionales de los complejos de kifunesina (columna 1), salacinol (columna 2) y ephotilone-A (columna 3) basadas en los factores R-NOE. Los puntos representan las conformaciones encontradas por los métodos de DRX (rojo), diferencias entre las intensidades experimentales y calculadas de STD-NMR (amarillo) y el campo de fuerzas OPLS2005 (verde). Los valores de las funciones han sido normalizados..... 103

Figura 5.1.8. Representación de los conformeros del complejo de ephotilone-A- β -tubulina obtenidos mediante el uso de técnicas de DRX (rojo), modelado molecular (verde), STD-NMR (amarillo), factores R-NOE (gris). (A) Solapamiento de las coordenadas de los conformeros de la ephotilone-A (B) Ampliación de la zona transversal de la cadena lateral de la ephotilone-A. (C) Datos de STD-NMR calculados a partir de las coordenadas de cada método de evaluación y los datos de STD-NMR experimentales (en negro). Los RMSD se obtienen tomando como referencia la estructura cristalográfica. 105

Figura 5.1.9. Capacidad de las funciones de ajuste de encontrar las estructuras de referencia para los complejos proteína-ligando entre varias poses de unión, considerando un amplio espectro de constantes químico-físicas. El eje de la

abscisa corresponde a la posición de la lista. La posición 1 es la considerada como la estructura con mejor ajuste. El eje de las ordenadas es el número de veces que se encontró alguna estructura de referencia sin importar la parametrización usada. 109

Figura 5.2.1. Diagrama de flujo del método implementado en el programa STD-MaPa (*STD Modelling Adjusting Parameter Approach*): Los óvalos representan el inicio y fin del programa, los romboides simbolizan cualquier tipo de entrada de datos, los rombos indican operaciones lógicas de respuestas binarias (sí o no) y los rectángulos simbolizan cualquier tipo de operación. Diferenciamos con fondo rosado, los elementos del método relacionados con la exploración paramétrica y con fondo verde, los elementos del método relacionados con la generación de las coordenadas del ligando en la unión..... 115

Figura 5.2.2. Constitución esquemática del programa STD-MaPa. En rosado, la interfaz accesible para el usuario; en naranja, módulos utilizadas por el gestor de programas; en morado, subrutinas; en verde, rutinas de terceros incluidas en el STD-MaPa. Las rutinas de terceros están bajo la Licencia Pública General (GNU) o el autor ha dado el consentimiento expreso para su implementación. STD-MaPa es un programa de uso libre y se distribuye bajo la licencia GNU. 119

Figura 5.2.3. Interfaz gráfica del STD-MaPa, en la que se muestra la ventana *Tabla* y el *Visualizador de Moléculas* (Jmol). 120

Figura 5.2.4. Interfaz gráfica del programa STD-MaPa dedicada al ajuste de las condiciones experimentales y parámetros químico-físicos. Se accede a ambas ventanas activando las pestañas de “*Molecular Info*” y “*NMR Info*” 122

Figura 5.2.5. (A) Estructuras cristalográficas generadas con el *Generador de Coordenadas* a partir de los compuestos de la base de datos DUD. Cada complejo es contado como una respuesta positiva si al menos se llega a generar una vez la estructura cristalográfica. Las respuestas afirmativas se

cuantificaron y normalizaron con respecto al total de complejos. (B) Representación esquemática de las dimensiones de los paralelepípedos en la β -tubulina (rojo). En azul, las dimensiones del espacio con el 100% de las estructuras cristalográficas generadas..... 126

Figura 5.2.6. Estructura química de los ligandos de referencia. (A) Kifunesina, (B) Salacinol, (C) Epotilona-A, (D) Podofilotoxina y (E) PM060327. 129

Figura 5.2.7. (A) Estructura cristalográfica del complejo de PM060327- T_2R . (B) Representación del sitio de unión con el modelo de la molécula de PM060327 localizada en la densidad electrónica. (C) Representación del sitio de unión con la estructura de la molécula de PM060327 unida a β 2-tubulina. 132

Figura 5.2.8. Datos de STD-NMR experimentales y calculados para los complejos de kifunesina-dGMII, salacinol-dGMII, epotilona-A- T_2R , podofilotoxina- T_2R y PM060327- T_2R . Los datos de STD-NMR experimentales están representados con líneas y círculos negros. Los datos de STD-NMR calculados están representados con líneas y cuadrados rojos. 134

Figura 5.2.9. Superposición de las poses generadas por el Autodock 4.2 en dos cajas cúbicas centradas en el sitio de unión y de aristas igual a 25 Å y 47 Å. (A) Kifunesina enlazada a dGMII. (B) Salacinol enlazado a dGMII. (C) Epotilona-A enlazada a la β -tubulina. (D) Podofilotoxina enlazada a la interfaz intradimérica entre α - y β -tubulina. (E) PM060327 enlazada a la β -tubulina. 138

Figura 5.2.10. Relaciones entre las estructuras correctas y el total generadas para cada complejo molecular por cada Factor R-NOE. El estudio se ha realizado con cinco complejos proteína-ligando y dos celdas de aristas de 25 Å y 47 Å. 141

Figura 5.2.11. Representación de las coordenadas de la unión obtenidas con el STD-MaPa. A la izquierda, las estructuras de referencia (gris) y las obtenidas con el STD-MaPa (verde) se encuentran superpuestas. A la derecha, los datos de STD-NMR experimentales (negro), los datos de STD-NMR calculados a

partir de las estructuras de referencia (rojo) y los datos arrojados por el STD-MaPa (naranja) son mostrados. 145

Figura 5.3.2. Esquema del método híbrido de resolución estructural de complejos moleculares basado en datos cristalográficos, TR-NOESY y STD-NMR. 158

Figura 5.3.3. (A) Representación de la estructura cristalográfica del complejo dactilolida- T_2R . La dactilolida no se representa porque no existe densidad electrónica suficiente para poder describir su estructura. (B) A la izquierda, representación de la densidad electrónica en $\beta 1$ - y $\beta 2$ -tubulina. En verde, el mapa de diferencia de contorno positivo de densidades electrónicas ($F_{obs} - F_{calc}$, φ_{calc}). En gris, la densidad electrónica observada (F_{obs} , φ_{calc}). A la derecha, superposición de la estructura de $\beta 1$ - y $\beta 2$ -tubulina del complejo de dactilolida- T_2R y la $\beta 2$ -tubulina del complejo de zampanolida- T_2R (PDB: 4I4T). En verde se representa el lazo-M, la zampanolida y la HIS229. (C) Superposición del complejo T_2R cristalizado con la dactilolida (en gris) y la zampanolida (azul claro) respectivamente. En la ilustración se procedió a aumentar los sitios de unión localizados en $\beta 1$ - y $\beta 2$ -tubulina. La ZMP está representada como esferas de Van der Waals amarillas y el lazo-M está resaltado con el color morado. 162

Figura 5.3.4. Caracterización mediante STD-NMR y TR-NOESY de la interacción dactilolida- α, β -tubulina. (A) Espectro unidimensional protónico y espectro de STD-NMR de la dactilolida en presencia de la α, β -tubulina. (B) Representación de la estructura 3D de la dactilolida y de las intensidades relativas de STD-NMR. Las esferas ubicadas en las coordenadas protónicas muestran las intensidades relativas de STD-NMR extraídas del espectro. Las esferas azules oscuro representan los protones sin datos de STD-NMR. (C) Espectro bidimensional de TR-NOESY. Las señales de los Tr-NOE son resaltadas. (D) Superposición de seis conformaciones de dactilolida que cumplen con las restricciones impuestas por las señales del espectro de TR-

NOESY. Las fuerzas de restricción utilizadas en la exploración conformacional están representadas con barras amarillas (50 kJ/mol\AA^2), naranjas (100 kJ/mol\AA^2) y rojas (200 kJ/mol\AA^2). 167

Figura 5.3.5. Superposición de las poses de unión determinadas por el STD-MaPa para el complejo de dactilolida y β -tubulina. Las dos primeras orientaciones de la lista del factor R_e son representadas. En la izquierda, representación del acoplamiento 1 ($R_e(\beta1)=0,29$) (azul claro) y 2 ($R_e(\beta2)=0,35$) (amarillo) en la $\beta1$ -tubulina del complejo T_2R . En la derecha, representación del acoplamiento 1 ($R_e(\beta1)=0,38$) (verde) y 2 ($R_e(\beta2)=0,46$) (marrón) en la $\beta2$ -tubulina del complejo T_2R 172

Figura 5.3.6. (A) Intensidades de STD-NMR del complejo de dactilolida- T_2R . Los datos experimentales se representan en negro, también se representan los datos calculados para el acoplamiento 1 en $\beta1$ (rojo) y $\beta2$ (azul claro), y el acoplamiento 2 en $\beta1$ (amarillo) y $\beta2$ (verde). (B) Refinamiento cristalográfico del acoplamiento 2 de la dactilolida y la $\beta2$ -tubulina. (C) (D). Superposición de los mapas de densidad electrónica (F_{obs} , φ_{calc}) y los mapas de diferencias ($2F_{obs}-F_{calc}$, φ_{calc}). En todos los casos con una resolución de $2,5 \text{ \AA}$. Los mapas de densidad electrónica están representados con una red azul. Los mapas de diferencia están representados en verde (contorno positivo, zona incluida en la estructura real y ausente en el modelo) y en rojo (contorno negativo, zona incluida en el modelo y ausente en la estructura real). 174

Figura 5.3.7. Diagrama de flujo del módulo externo al STD-MaPa encargado de resolver la estructura de los complejos moleculares con múltiples modos de unión. El STD-MaPa genera las coordenadas de los múltiples acoplamientos moleculares y optimiza los parámetros químico-físicos. En naranja se identifican las acciones realizadas por el módulo externo al STD-MaPa. En verde se identifican las acciones manuales y externas al programa. 179

Figura 5.3.8. (A) Representación de los dos conformeros que minimizan el valor del factor R_e para el modo de unión mayoritario en $\beta1$ -tubulina. (B) Representación de los dos conformeros que minimizan el valor del factor R_e

para el modo de unión minoritario en $\beta 1$ -tubulina. (C) Distribución de intensidades de STD-NMR de los datos experimentales (negro) y los modos de unión resueltos por el STD-MaPa (rojo y naranja). Los colores de los datos corresponden con el par de estructuras representadas a la izquierda. (D) Representación de los modos de unión mayoritarios en presencia de los átomos de oxígenos de aguas resueltos para la estructura cristalográfica. ... 181

Figura 6.1.1 Variación de los cuatro términos de $\delta R_e/\delta S_e$ en función de la diferencia entre el valor experimental y el teórico. En el cálculo se consideró un ligando hipotético con 24 protones con la misma saturación experimental ($S_{exp}=12$). Los valores de la saturación teórica (S_{cal}) se fijaron con valores crecientes de saturación que van desde $S_{cal}=1$ hasta $S_{cal}=24$ 190

Figura 6.2.1 Representación esquemática de la redistribución de la saturación efectuada mediante los términos ΔM y ΔM_w en presencia de desviaciones en la saturación experimental debida a artefactos experimentales. Las líneas representan la media aritmética y ponderada de las intensidades..... 199

Figura 6.2.2 Patrones de STD-NMR experimentales (líneas sólidas con círculos) y calculados (líneas discontinuas con cuadrados) a partir de la estructura resuelta por el programa STD-MaPa en diferentes tiempos de saturación experimentales. Se representaron los patrones de STD-NMR en gris ($t_{sat}=0,50$ s), azul ($t_{sat}=0,75$ s), verde ($t_{sat}=1,00$ s), rojo purpúreo ($t_{sat}=1,50$ s), violeta ($t_{sat}=2,00$ s), naranja ($t_{sat}=2,50$ s) y verde olivo ($t_{sat}=3,00$ s). En la esquina inferior derecha se encuentra la estructura química de la dimanosa.207

Figura 6.4.1. Representación del mecanismo de reacción covalente de la dactilolida y la zampanolida en el sitio interno de la β -tubulina. El mecanismo de reacción covalente está representado en tres pasos que se diferencian con el naranja (estructuras antes de la reacción covalente), negro (estructura intermedia) y verde (estructura posterior a la reacción covalente) (A) Superposición de la estructura de $\beta 1$ - y $\beta 2$ - tubulina en presencia de DAC (naranja) y ZMP (verde) respectivamente. (B) Representación, según la Teoría

de Enlace de Valencia (TEV), del mecanismo de la reacción covalente entre el macrociclo de ZMP o DAC con el residuo de histidina 229. En naranja, se representan los pasos que están relacionados con la estructura del complejo de DAC-T₂R obtenida mediante el método híbrido. En verde, se representan los pasos que están relacionados con la estructura del complejo de ZMP-T₂R obtenida mediante DRX. En negro, el paso intermedio del mecanismo de unión covalente. (C y D) Cambio de perspectiva de la imagen “A” mediante un giro de 90° con respecto un eje horizontal. (E y F) Cambio de perspectiva de las imágenes “C” y “D” respectivamente mediante un giro de -90° con respecto a un eje vertical. 217

Figura 6.5.1 . Representación de los modos de unión de la podofilotoxina y sus dos análogos 2 y 3. La podofilotoxina se representó en naranja y sus derivados en gris. (A) Superposición de la estructura cristalográfica de la podofilotoxina PDB 1SA1 (azul claro) y la estructura obtenida con el STD-MaPa (naranja). (B) Superposición de la estructura de la podofilotoxina y el compuesto 2. (C) Superposición de la estructura de la podofilotoxina y el compuesto 3 (conformación A de la Figure 8 de (Antúñez-Mojica, Rodríguez-Salarichs et al. 2016)). En el inferior de la imagen, se encuentra la comparación entre los patrones de STD-NMR experimentales y los calculados a partir de los modelos moleculares obtenidos con el STD-MaPa de los compuestos 2 y 3..... 224

Figura 6.5.2 Representación de las estructuras químicas de la epotilona-A y la azatilonona-2. Las diferencias que presentan en sus estructuras químicas están representadas en rojo y azul (A) Representación de la azatilonona-2 unida al sitio interno de la β -tubulina. Las coordenadas fueron obtenidas con un primer ciclo con el STD-MaPa (B) Superficie del factor R-NOE obtenida a partir de la exploración conformacional con el Macromodel y el STD-MaPa. 144 conformaciones de la azatilonona-2 fueron obtenidas mediante la rotación de los enlaces C15-C16 y O-C(CH₃)₃. Los enlaces que se rotaron están resaltados en amarillo en la imagen “A” (C) Patrón de STD-NMR experimental (negro) y teórico (rojo) del conformero A; (D) Patrón de STD-NMR experimental (negro) y teórico (rojo) del conformero B; (E) Modelo refinado del conformero A unido a la β -tubulina; (F) Modelo refinado del conformero B unido a la β -tubulina. (G)

Superposición de las estructuras del conformero A (naranja) y B (gris); (**H**)
 Superposición del modelo 3D del complejo azatilon-2- β -tubulina (conformero
 B en naranja) y el complejo epotilon-A- β -tubulina (azul claro). 227

Figura 8.2.1. En la parte superior, se representan en los gráficos los valores
 máximos de las desviaciones de los factores R-NOE, al variar los parámetros
 químico-físicos para los tres complejos moleculares. En la parte inferior
 aparece el estudio estadístico de igualdad de varianza de Levene
 ([http://scistatcalc.blogspot.com.es/2013/12/levenes-test-equality-of-](http://scistatcalc.blogspot.com.es/2013/12/levenes-test-equality-of-variances.html)
[variances.html](http://scistatcalc.blogspot.com.es/2013/12/levenes-test-equality-of-variances.html)), el de Barlett
 (<http://scistatcalc.blogspot.com.es/2013/10/bartletts-test-calculator.html>) y la
 prueba de normalidad Shapiro-Wilk realizada con el programa QtiPlot. En la
 izquierda se encuentran las variaciones máximas de los factores R-NOE para
 el sistema con kifunesina, salacinol y epotilon-A al modificar los parámetros de
 la matriz cinética (**K**). En la izquierda se encuentran las variaciones máximas
 de los factores R-NOE para el sistema con kifunesina, salacinol y epotilon-A al
 modificar los parámetros de la matriz de relajación (**R**). 245

Índice de tablas

Tabla 5.1.1. Conjunto de parámetros que minimizan el factor R_e en los complejos de kifunesina-dGMII, salacinol-dGMII y epotilona-A- β -tubulina.....	73
Tabla 5.1.2. Resultado del estudio estadístico ANOVA de una cola.	90
Tabla 5.1.3. Prueba estadística post hoc de Tukey.....	93
Tabla 5.1.4 Desviación cuadrática media de los confórmeros descritos por la DRX, RMN y el modelado molecular.....	100
Tabla 5.1.5. Porcentaje de acierto de las funciones de ajuste en el reconocimiento de estructuras de referencia entre 300 estructuras generadas por el AutoDock 4.2.....	110
Tabla 5.2.1. Tabla de datos cristalográficos del complejo PM060327-T ₂ R. ...	130
Tabla 5.2.2. Rangos paramétricos explorados en los cinco sistemas de referencia	140
Tabla 5.3.1. Tabla de datos cristalográficos del complejo dactilolida-T ₂ R.....	163
Tabla 5.3.2. Resultados del STD-MaPa de las dos primeras posiciones de las listas creadas por el factor R_e en β 1- y β 2- tubulina.	171
Tabla 5.3.3. Parámetros y concentraciones resueltas a partir de las ecuaciones EGC-CORCEMA.....	177
Tabla 6.2.1 Ajuste paramétrico realizado por el STD-MaPa con el factor R_e al resolver la estructura del complejo dimanosa-2G12 en diferentes tiempos de saturación.....	204
Tabla 6.2.2 Ajuste paramétrico realizado por el STD-MaPa con el factor R_t al resolver la estructura del complejo dimanosa-2G12 en diferentes tiempos de saturación.....	205

Tabla 6.3.1. Conjunto de parámetros que minimizan los factores R_e y R_t en los complejos cristalográficos de podofilotoxina- α,β -tubulina, PM060327- α,β -tubulina y epotilona-A- β -tubulina.	212
---	-----

Abreviaturas

NOE	<i>El efecto nuclear Overhauser</i>
STD-NMR	<i>Experimento de Diferencia de Transferencia de Saturación</i>
TR-NOESY	<i>Experimento de NOE Transferido</i>
Tr-NOE	<i>NOE transferido</i>
dGMII	<i>Proteína α-manosidasa II de la <i>Drosophila melanogaster</i></i>
k_{on}	<i>Constante de velocidad de asociación</i>
k_{off}	<i>Constante de velocidad de disociación</i>
Taxol®	<i>Nombre comercial del fármaco Paclitaxel</i>
ANOVA	<i>Prueba estadística basada en el análisis de la varianza</i>
RMSD	<i>Desviación cuadrática media</i>
SNR	<i>Relación señal-ruido</i>
Crio-ME	<i>Criomicroscopía electrónica</i>
DRX	<i>Difracción de Rayos X</i>
RMN	<i>Espectroscopia de Resonancia Magnética Nuclear</i>
PC	<i>Ordenador Personal</i>
DUD	<i>Base de datos de Directorio de Señuelos Útiles</i>
RAM	<i>Memoria de acceso aleatorio</i>
MSA	<i>Agente estabilizante de microtúbulo</i>

PEG	<i>Polietilenglicol</i>
DAC	<i>Macrólico (-)-dactilolida</i>
EGC-CORCEMA	<i>Ecuaciones Generales de Concentraciones del CORCEMA</i>
CD	<i>Dicroísmo Circular</i>
FCS	<i>Espectroscopia de Correlación de Fluorescencia de dos fotones</i>
GndCl	<i>Cloruro de guanidinio</i>
TEV	<i>Teoría de Enlace de Valencia</i>

Resumen

Resumen

El conocimiento estructural de los complejos moleculares entre proteínas y moléculas pequeñas permite obtener información útil para el estudio de la maquinaria celular. En particular, la resolución estructural de complejos proteína-ligando de baja afinidad tiene vital importancia, por su presencia en gran parte de los mecanismos de regulación celular. En este último caso, la resolución estructural es ardua y en muchos casos llega a ser inviable debido a sus bajas constantes de asociación. En estas condiciones, es difícil extraer información estructural suficiente por DRX o la crio-ME para poder resolver analíticamente la estructura de los complejos. Sin embargo, diversas técnicas de RMN permiten extraer parte de la información estructural de este tipo de reacciones.

En un intento de resolver los complejos moleculares bajo la influencia de un intercambio químico rápido, se diseñó una metodología híbrida basada en datos de RMN y DRX dedicada a la resolución estructural de complejos proteína-ligando con bajas constantes de asociación. La metodología combina la estructura cristalográfica de la proteína, la conformación basada en los Tr-NOEs y la orientación basada en datos de STD-NMR para resolver la estructura de los complejos proteína-ligando de baja afinidad.

Para implementar los datos de STD-NMR con el resto de las técnicas es necesario el desarrollo de una metodología fiable que determine automáticamente la orientación de la unión del ligando en el sitio de unión. Por lo que, se realizó el diseño de nuevas variantes de la función de error residual usada para evaluar la calidad de las estructuras obtenida mediante datos de RMN (factor R-NOE) con el fin de mejorar su robustez. Las nuevas variantes del factor junto con el factor R-NOE se implementaron en el programa STD-MaPa. STD-MaPa es el núcleo de la metodología híbrida de RMN-DRX y fue creado para automatizar el acoplamiento guiado por datos de STD-NMR de las

moléculas de ligando en el sitio activo de los complejos proteína-ligando. La metodología implementada en el programa fue probada usando como referencia cinco complejos moleculares que tienen sus estructuras resueltas por DRX. El STD-MaPa resolvió correctamente las cinco estructuras cristalográficas.

Finalmente, la estructura de la dactilolida unida al tetrámero de α,β -tubulina fue estudiada. La dactilolida forma parte de una nueva familia de drogas capaces de unirse covalentemente a la β -tubulina evitando así la resistencia tumoral por sobreexpresión de bombas de membrana. Esta droga se une covalentemente a la proteína pero presenta una cinética de asociación muy lenta, lo que impide su resolución estructural por DRX.

En nuestro trabajo se llegó a resolver la estructura de la dactilolida unida no covalentemente al tetrámero de α,β -tubulina con la metodología híbrida RMN-DRX. Además, se identificaron dos orientaciones en la subunidad de β -tubulina menos accesible del cristal del tetrámero T₂R en una proporción del 83% y 17%. Al analizar la orientación predominante y la estructura cristalográfica del homólogo estructural zampanolida se describió el mecanismo de reacción covalente de esta familia de droga en la unión a la β -tubulina. Mecanismo que permitirá el diseño de nuevos fármacos que presenten la unión covalente y eludan los mecanismos de resistencia a fármacos de las células cancerígenas.

Abstract

Molecular recognition plays a key role in many cellular processes. Knowledge about the structure of the molecular complexes implicated in these cellular events is therefore important to understanding cellular mechanisms. For cell function, it is important that most of the processes are reversible, and thus the molecules are often weakly associated.

X-ray diffraction is a widely used technique in the structural analysis of molecular complexes; however, it has the disadvantage of not being useful in solving the structure of molecular complexes with low association constants. In these complexes, symmetry in the crystal is affected by the abundance of empty binding pockets. In the case of reversible molecular complexes formed by ligands with low molecular weight with macromolecules, the crystallographic methods are not able to extract information in order to describe the entire structure. This disadvantage can be resolved using complementary Nuclear Magnetic Resonance (NMR) techniques, and integrating the structural information obtained by both techniques.

NMR is an extremely versatile technique that can be used to obtain structural information from molecules or their molecular complexes. For biological systems, NMR is a very useful technique to study the structure of proteins, lipids or polysaccharides. Two-dimensional transferred nuclear Overhauser effect spectroscopy (TR-NOESY) and Saturation Transfer Difference (STD-NMR) are NMR methods widely used to study biological interactions. Both techniques allow studying molecular complexes with low association constants because they can measure magnetization from molecules in solution. TR-NOESY provides information about the conformation of low molecular weight ligands when bound to macromolecules and STD-NMR contains information about the binding epitope of a small molecule bound to a protein.

In this work, we have designed a hybrid methodology, able to obtain the structures of macromolecular complexes by using STD-NMR data, TR-NOESY conformer and a partially solved structure obtained by X-ray diffraction. In order to ensure the methodology resolves the molecular complexes analytically, a new auto-consistent algorithm is implemented to solve the binding pose of protein-ligand complexes based on STD-NMR data. For ensuring that the algorithm is auto-consistent, five new descriptors have been implemented to differentially evaluate the fitting of the STD-NMR maps back calculate from a binding pose with the experimental data.

The algorithm uses the structure of macromolecule partially resolved by X-ray diffraction, STD-NMR data and TR-NOESY conformation of ligand to estimate chemical-physical parameters of the system (such as the apparent binding constant and the protein apparent correlation time) and to solve the structure of molecular complex fitting the experimental STD-NMR map and the calculated one. The structure of the molecular complex with the new structural information obtained can then be iteratively refined using steps of electron density and STD-NMR maps fitting until the solved structure is found to be consistent with both techniques.

Finally, the structure of dactylolide, a microtubule-stabilizing agent, was solved bound to the luminal taxane site in T₂R tetramers using the hybrid methodology. Dactylolide has a fast chemical interchange between free and bound states, which leads to a reduction of crystallographic resolution in the binding site. But, There was applied the hybrid methodology and structure of dactylolide was solved bound to the β -tubulin subunit deeper in the crystal. The analysis of dactylolide-T₂R structure and the crystallographic structure of zemanolide, a structural homologous of dactylolide, allows characterizing the covalent reaction mechanism of drugs in the Taxol site.

Introducción

1 Introducción

El estudio estructural de los complejos biológicos nos permite acceder a valiosa información sobre sus propiedades químico-físicas. En el contexto biológico, la caracterización estructural de las moléculas permite describir la organización y funcionamiento de sistemas tan complejos como las células. La célula es la unidad morfológica y funcional de todo ser vivo; está compuesta por diversos compuestos moleculares como las proteínas, polisacáridos o nucleótidos. La interrelación entre todos estos componentes controla la vida celular. Por tanto, conocer la estructura y mecanismos celulares de todos los compuestos moleculares es uno de los objetivos fundamentales de la ciencia moderna.

Las técnicas de determinación estructural aplicadas al estudio de las macromoléculas son diversas. El desarrollo y renovación que han experimentado durante años ha permitido potenciar sus capacidades y aplicaciones. Múltiples ejemplos de este desarrollo se pueden encontrar en las últimas dos décadas, tiempo en el que se han llegado a resolver estructuras de una alta complejidad. Entre todos los antecedentes que encontramos, haremos alusión a solo dos que fueron notables por su impacto en el estudio estructural y mecanicista de las celulares:

La elucidación estructural del motor molecular ATP sintasa con una resolución de 3,9 Å (Stock, Leslie et al. 1999, Stock, Gibbons et al. 2000); y la visualización *in-vitro* y en tiempo real de la dinámica de la proteína myosin V sobre un filamento de actina (Ando et al., 2013; Koder and Ando, 2014). Las técnicas que hicieron posibles ambos estudios, son fieles exposiciones de los avances tecnológicos acaecidos hasta la fecha. En la actualidad, tanto los protocolos como las herramientas experimentales siguen evolucionando, y con sus avances es más común el derrumbamiento de viejas barreras.

Como ejemplos de técnicas estructurales empleadas en la biología se encuentran las técnicas de Difracción de Rayos-X (DRX) y Resonancia

Magnética Nuclear (RMN). La descripción molecular de muchos procesos celulares ha sido posible gracias a los avances metodológicos en DRX y RMN. Aunque sus aportes han sido inmensurables, sus limitaciones pueden derivar, en ciertos casos, hacia la incorrecta descripción de la estructura y la obtención de información sin utilidad biológica.

1.1 Difracción de Rayos-X

La Difracción de Rayos-X (DRX) es una técnica experimental que permite caracterizar la estructura atómica de las biomoléculas. Las bases de su fundamento físico son similares a las de cualquier otra técnica que utilice radiación electromagnética. En la DRX se hace incidir una fuente de rayos-X sobre un sólido cristalino que absorbe y dispersa la radiación. La interacción entre el haz incidente y la nube electrónica del material dispersa la radiación electromagnética (Figura 1.1.1-A). La dependencia entre la dispersión y las características físicas del material, tales como el orden, composición y estructura, da lugar a un patrón de difracción único que posteriormente se procesa para obtener la estructura atómica del material. Este fenómeno se produce a consecuencia de las restricciones angulares impuestas por el orden molecular existente en un sólido cristalino. Además, sólo ocurre cuando la amplitud de la longitud de onda de la radiación incidente coincide en tamaño con la distancia entre planos atómicos de un cristal (Bragg and Bragg 1913) (Figura 1.1.1-B). Esta coincidencia permite caracterizar las estructuras moleculares a nivel atómico.

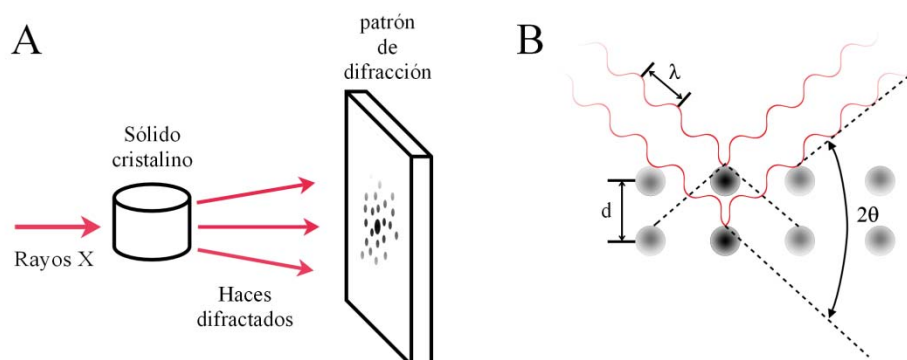


Figura 1.1.1 Representación del fenómeno de difracción de rayos-X en un sólido cristalino (A). Corrimiento de fase que da lugar a las interferencias constructivas entre las ondas según la ley de Bragg (B).

Las intensidades de difracción que se miden en un experimento de DRX provienen de la interferencia constructiva de la radiación dispersada por la nube electrónica de los átomos cuando están en las posiciones de Bragg (Bragg and Bragg 1913). Si se desea encontrar las posiciones de los átomos que dan lugar al patrón de difracción, se tiene que analizar la función de densidad electrónica ($\rho(\mathbf{r})$) del material; ya que los máximos de la función coinciden con las posiciones de los átomos. Describir la función $\rho(\mathbf{r})$ a partir de las intensidades de DRX es un proceso complicado porque depende de la transformada de Fourier inversa del factor de estructura $\mathbf{F}(\mathbf{H})$.

$$\rho(\mathbf{r}) = \frac{1}{V} \sum_{\mathbf{H}} \mathbf{F}(\mathbf{H}) e^{-2\pi i \mathbf{H} \cdot \mathbf{r}} \quad 1.1.1$$

El factor de estructura $\mathbf{F}(\mathbf{H})$ es una función compleja de variable real que se relaciona con las intensidades de difracción de la siguiente manera:

$$I(\mathbf{H}) = k \times I_0 \times L \times P \times A \times |\mathbf{F}(\mathbf{H})|^2 \quad 1.1.2$$

k es el factor de escala, I_0 es la intensidad del haz incidente; L y P son los factores de Lorentz y polarización respectivamente; y A es el factor de absorción y extinción. Por lo tanto, a partir de las intensidades extraídas del patrón de difracción, sólo podemos obtener experimentalmente los módulos del factor de estructura ($|\mathbf{F}(\mathbf{H})|$), debido a que $\mathbf{F}(\mathbf{H})$ es un número complejo.

$$\mathbf{F}(\mathbf{H}) = \sum_j^N f_j e^{(2\pi i \mathbf{H} \cdot \mathbf{r}_j)} \quad 1.1.3$$

$$\mathbf{F}(\mathbf{H}) = F(\mathbf{H}) e^{(i\alpha(\mathbf{H}))} \quad 1.1.4$$

Este problema descrito es conocido como el problema de las fases, y su resolución le supuso la concesión del premio Nobel de Química en 1985 a Karle y Hauptman (Hendrickson 1986).

Los primeros resultados relevantes en sistemas biológicos se dieron a conocer a partir de los años 50. Por primera vez, en 1953 se resolvió el modelo de doble hélice del ADN a partir de datos de difracción (Watson et al., 1953). Una década más tarde, al inicio de 1960, se determinó la estructura de la hemoglobina y se llegó a obtener la sorprendente resolución de 5.5 Å para dicha molécula (Perutz et al., 1960).

Los avances en la cristalización de macromoléculas biológicas han permitido la obtención de monocristales con una alta simetría y sin imperfecciones. Además, los difractómetros de monocristales son cada vez más potentes y accesibles, lo que permite resolver estructuras de biomoléculas de gran tamaño

y complejidad. En la actualidad, la difracción de rayos-X sigue siendo la técnica preferida cuando se desea predecir la estructura de las proteínas. Un factor clave en el impulso del uso de la DRX fue la creación de bases de datos como por ejemplo el Banco de Datos de Proteínas (o PDB, por su acrónimo en inglés). La aparición de grandes bases de datos impulsa la creación de nuevos algoritmos de resolución estructural (McCoy, Grosse-Kunstleve et al. 2007, Adams, Afonine et al. 2010) que reducen el tiempo y el esfuerzo del proceso resolutivo. Las nuevas metodologías tienen implementado el método de Reemplazo Molecular (*"Molecular Replacement"*) para resolver el problema de las fases (Drenth 2013). En este método se trata de buscar el modelo que mejor ajusta las intensidades experimentales del compuesto desconocido a partir de un compuesto similar con su estructura ya resuelta.

Pese a los avances en la DRX, esta técnica sigue teniendo limitaciones como cualquier otra técnica basada en la radiación electromagnética. Una de las desventajas se encuentra en el proceso de cristalización; éste es engorroso, y no asegura la formación de monocristales para todos los sistemas biológicos. Además, la estructura cristalina del complejo biológico no es necesariamente idéntica a la estructura en solución, por lo que la estructura resuelta puede no ser la biológicamente activa. Otra desventaja que nos encontramos está relacionada con la resolución estructural. En las estructuras de las proteínas es muy común encontrarse regiones flexibles donde la movilidad en el cristal reduce la simetría: en estas regiones la resolución disminuye al extremo que en ocasiones no se logra describir. Por último, si deseamos resolver la estructura de un complejo molecular por DRX, éste tiene que asociarse fuertemente, para así favorecer la unión del ligando al receptor y reducir los sitios de unión vacíos.

1.2 Resonancia Magnética Nuclear

La Espectroscopia de Resonancia Magnética Nuclear (RMN) es una herramienta muy versátil y con ella se puede extraer información tanto estructural como dinámica de los complejos biológicos en solución. Para los sistemas biológicos, es de gran utilidad para el estudio de moléculas pequeñas, proteínas, lípidos y polisacáridos. La mayor ventaja que nos proporciona esta técnica radica en la posibilidad de estudiar las propiedades estructurales, dinámicas y funcionales de las biomoléculas en solución, bajo condiciones parecidas a las nativas.

La Resonancia Magnética Nuclear fue desarrollada en 1945 por los científicos americanos Felix Bloch (1905-1983) y Edward M. Purcell (1912-1997); hecho que les permitió obtener en 1952 el Premio Nobel de Física. En los primeros años, la técnica demostró su potencial en la determinación estructural de moléculas orgánicas aunque en aquellos momentos el estudio de moléculas de gran complejidad quedaba excluido debido a la pobre resolución que se obtenía. Sin embargo, en la segunda mitad de la década de los '60, la técnica fue impulsada por la implementación de la modalidad de la Transformada de Fourier (TF) (Ernst 1966, Ernst and Anderson 1966, Richard 1969).

La TF permitió aumentar significativamente la resolución espectral de la técnica, hecho que posibilitó el estudio de sistemas más complejos (Ernst and Anderson 1966). Las nuevas y mejoradas ventajas adquiridas con este método se unieron a la creciente evolución de los equipamientos y metodologías experimentales. Aunque son muchos los aportes que han permitido el desarrollo de la RMN, entre los más significativos se encuentran la aparición de imanes superconductores, el creciente empleo de la resonancia de heteronúcleos y la introducción de la espectroscopia multidimensional (Markley and Opella 1997). Todos ellos juntos llegaron a permitir el estudio de macromoléculas de interés para la biología mediante la RMN.

En los experimentos de RMN se extrae información sobre la interacción de las partículas nucleares (protones y neutrones), en presencia de diferentes campos magnéticos. En el núcleo atómico, cada protón o neutrón se puede aparear con una partícula de su misma clase con espín antiparalelo. El espín es una propiedad intrínseca que presentan los electrones, neutrones y protones, y se define como el valor del momento angular que tiene asociado cada una de las partículas que forman los átomos. El espín de una partícula puede describir varios estados cuantizados de energías y cada uno de ellos se caracteriza por tener un número cuántico de espín (m), para el núcleo de ^1H puede adoptar los valores $+\frac{1}{2}$ o $-\frac{1}{2}$.

Si las partículas nucleares se encuentra apareadas (neutrones o protones), el espín neto es igual a cero ($I=0$). Pero, cuando los núcleos tienen números de protones y neutrones impares, dan lugar a un espín neto mayor que cero ($I>0$). En estos casos, el núcleo va a tener un momento magnético asociado (μ) que depende de la dirección del espín.

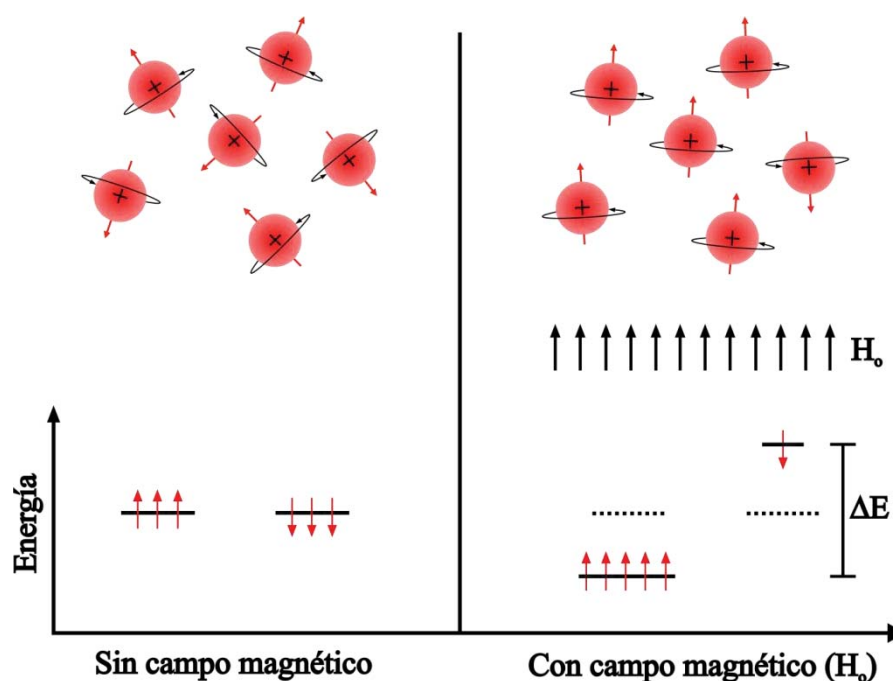


Figura 1.2.1 Representación de la situación de equilibrio de los espines nucleares con $I=1/2$ en ausencia de un campo magnético externo (izquierda) y cuando están bajo la influencia de un campo magnético H_0 (derecha). En la zona inferior de la figura se muestra los niveles de energía en ausencia y presencia de H_0 respectivamente.

En ausencia de un campo magnético y en situación de equilibrio, encontraremos el mismo número de átomos en cada uno de los estados de espín. En estos casos, los niveles de energía de los momentos angulares se encontrarán degenerados (Figura 1.2.1, izquierda). Sin embargo, en presencia de un campo magnético, uno de los ejes de los momentos angulares va a coincidir con la dirección del campo. Como consecuencia, el estado que se encuentre alineado al campo va a tener una menor energía que el estado con el espín en dirección opuesta (Figura 1.2.1, derecha). La diferencia de energía entre los estados es proporcional a la fuerza del campo magnético externo y se define por la siguiente ecuación,

$$\Delta E = h\nu = h \frac{\omega_0}{2\pi} = h \frac{\gamma}{2\pi} B_0 \quad 1.2.1$$

donde h es la constante de Planck, ν es la frecuencia de resonancia, ω_0 es la frecuencia de resonancia de Larmor, γ es la constante magnetogírica de la partícula y B_0 es el campo magnético externo.

El procedimiento a seguir en un experimento de RMN es relativamente simple. Primero, se debe tener un campo magnético intenso, homogéneo e invariante en el tiempo (B_0). En el transcurso del experimento se aplicará una serie de pulsos de radiofrecuencias (B_1) en dirección perpendicular al campo B_0 . La aplicación de los pulsos de radiofrecuencia está dirigida a distorsionar temporalmente el estado de equilibrio obtenido en presencia del campo B_0 . Por eso, los valores de frecuencia de los pulsos tienen que ser equivalentes a la diferencia de energía entre niveles de espín, para poder inducir la transición entre los estados de espín. Al finalizar cada pulso, el sistema va a tender espontáneamente al equilibrio mediante la emisión del exceso de energía en forma de una onda de radio a la frecuencia de Larmor de los espines. Esta señal se amplifica y digitaliza convenientemente para poder obtener en un procesamiento posterior un espectro de RMN. La relajación, al igual que la magnetización, depende del tipo de átomo y del entorno químico. Por ambos motivos, a partir del procedimiento descrito recientemente se puede extraer información estructural.

A causa de la interacción física entre los momentos angulares de los espines, éstos dejan de ser independientes y acoplan sus momentos magnéticos. Para explicar este fenómeno se puede considerar, desde un punto de vista simple, que los espines son pequeños imanes que poseen momentos magnéticos. Bajo esta suposición, el acoplamiento ocurriría igualmente mediante la interacción de los campos magnéticos entre dos imanes (Figura 1.2.2). Por tanto, el espín nuclear se acoplará con los espines electrónicos a través de la interacción de

sus respectivos campos magnéticos, para así transmitirse a través del enlace y afectar el momento magnético del núcleo vecino.

Modelo vectorial de Dirac

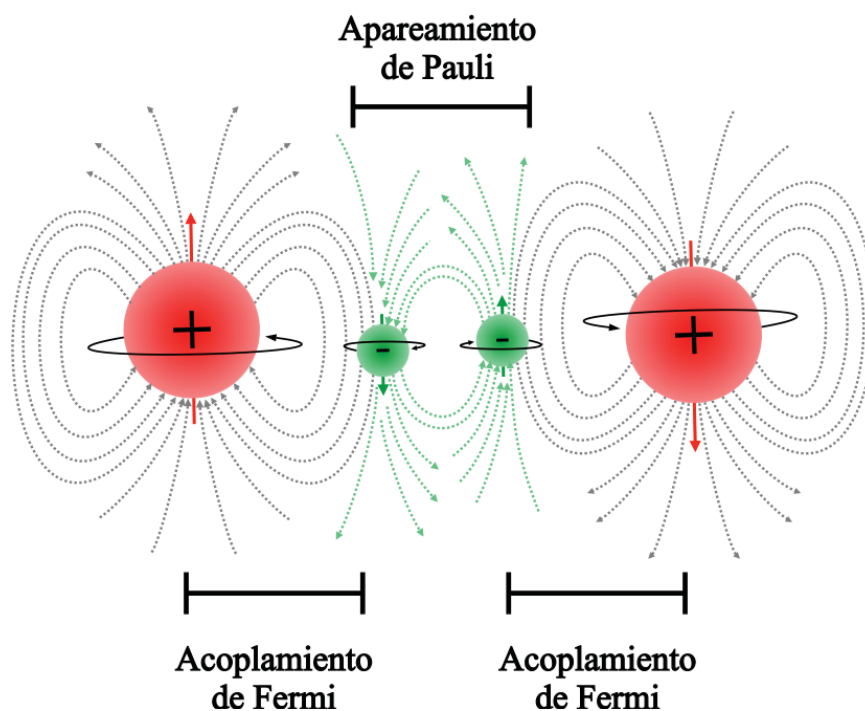


Figura 1.2.2 Modelo vectorial de Dirac y los diferentes acoplamientos entre los espines de las partículas atómicas.

El mecanismo del acoplamiento espín-espín que se acaba de describir se llama acoplamiento indirecto de espín, ya que tiene lugar a través de los electrones del enlace. Por consiguiente, el acoplamiento directo va a tener lugar a través del espacio entre dos núcleos. En este caso, los núcleos se comportan como dos dipolos que en ciertos casos darán lugar al denominado Efecto Nuclear Overhauser.

1.2.1 Efecto Nuclear Overhauser

El Efecto Nuclear Overhauser o efecto NOE (por su acrónimo en inglés) es un mecanismo de relajación que permite transmitir polarización entre espines nucleares acoplados dipolarmente.. Como cualquier interacción dipolar, el efecto NOE depende de la distancia entre los espines, y por este motivo su magnitud se aprovecha para extraer información estructural de los sistemas moleculares. En la actualidad, el efecto NOE es muy usado en la determinación conformacional de moléculas orgánicas y macromoléculas.

El efecto Overhauser fue descrito y demostrado experimentalmente en 1953 (Carver and Slichter 1953, Overhauser 1953). Pero no fue hasta 1955, con la aparición de las ecuaciones de Solomon, cuando se describió su fundamento con una base teórica (Solomon 1955). Solomon llegó a describir matemáticamente el proceso de relajación entre dos espines aislados a partir de las probabilidades de transición de cero (w_0), uno (w_1 y w_1') y doble (w_2) cuanto entre cuatro estados de espines (Figura 1.2.3).

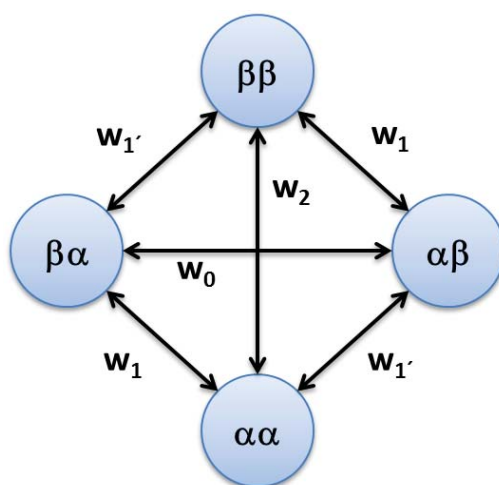


Figura 1.2.3 Diagrama de transiciones de dos espines nucleares aislados entre los cuatro estados de espines posibles, cuando se considera un sistema de núcleos con $I=1/2$, no acoplados escalarmente y en ausencia de otro tipo de interacción que no sea dipolar. w_i para $i=0,1,2$ es la probabilidad de transición entre estados.

En este sistema de espines aparecen dos nuevas transiciones (w_0 y w_2) que no se encuentran en un sistema de un solo espín; siendo además las responsables de la transferencia de la magnetización entre espines:

$$\sigma_{ij} = w_2(ij) - w_0(ij); i \neq j \quad 1.2.2$$

donde σ_{ij} es la constante de relajación cruzada y describe la velocidad de transición dipolo-dipolo entre el espín i y j que da lugar al efecto NOE. Por otro lado, los procesos responsables de restablecer el estado de equilibrio fueron descritos con la constante de velocidad de la relajación longitudinal (ρ_{ij}).

$$\rho_{ij} = w_0(ij) + 2w_1(ij) + w_2(ij); i \neq j \quad 1.2.3$$

$$\rho_{ij}' = w_0(ij) + 2w_1'(ij) + w_2(ij); i \neq j \quad 1.2.4$$

Para llegar a describir el efecto NOE, las tres ecuaciones (1.2.2, 1.2.3 y 1.2.4) fueron añadidas a las ecuaciones macroscópicas de Bloch (Bloch 1946) adaptadas a un sistema de dos espines. En su resolución consideraron el estado estacionario y la presencia de uno de los espines saturado (Ecuación 22 en (Solomon 1955)).

Finalmente, el efecto NOE sobre el espín i (η_i) se describe de la siguiente manera.

$$\eta_i = \frac{I_i - I_i^0}{I_i} = \frac{w_2(ij) - w_0(ij)}{w_0(ij) + 2w_1(ij) + w_2(ij)}; i \neq j \quad 1.2.5$$

Inicialmente se empleaban experimentos de NOE monodimensional que permitían abordar la caracterización de la estructura de compuestos orgánicos pequeños. Posteriormente con el desarrollo de la espectroscopía de RMN bidimensional se pudieron adquirir experimentos 2D NOESY que permitieron resolver estructuras mucho más complejas (Jeener, Meier et al. 1979, Meier and Ernst 1979). Al inicio de los '80, gracias a la espectroscopia bidimensional y a la creación de algoritmos geométricos de determinación estructural, se llegó a obtener la primera estructura espacial de una proteína basada en RMN (Kumar, Ernst et al. 1980, Wüthrich, Wider et al. 1982, Havel and Wüthrich 1984).

1.2.2 Espectroscopia bidimensional

La espectroscopia bidimensional está formada por un grupo de métodos de RMN a partir de los cuales se obtiene un espectro definido por dos ejes de frecuencias correlacionadas. La idea revolucionaria de la espectroscopia bidimensional fue introducida por el físico belga Jean Jeener en la escuela de verano “Ampere International Summer School” celebrada en 1971 (Goldman 1994). Este profesor propuso una secuencia de dos pulsos simples que producen, posterior a una transformación de Fourier bidimensional, un espectro en dos dimensiones (Figura 1.2.4). Su trabajo sobre el tema nunca fue publicado, pero la idea fue bien acogida por Richard R. Ernst que la desarrolló en los subsiguientes años.

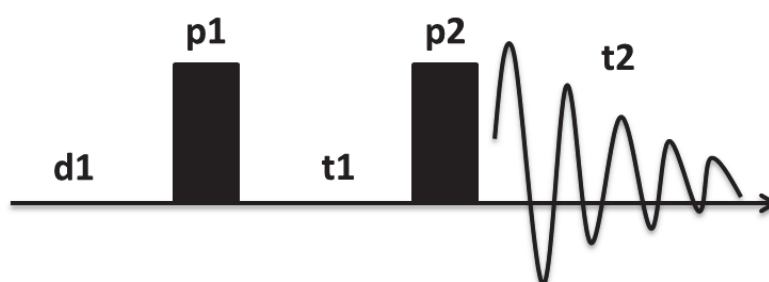


Figura 1.2.4 Espectroscopia de correlación. Consiste en un simple pulso de radiofrecuencia (p1), seguido por un tiempo de evolución (t1) y un posterior pulso de radiofrecuencia (p2). Después del último pulso se inicia el período de medida (t2).

El experimento nacido de la secuencia de pulsos propuesta por el profesor Jean Jeener se denominó espectroscopia de correlación (COSY) (Maudsley, Moller et al. 1977, Nagayama, Wüthrich et al. 1977). La secuencia COSY está diseñada para obtener información sobre las conexiones intramoleculares a través de los enlaces (acoplamiento escalar). Esta circunstancia se logra a través de una familia de pulsos de radiofrecuencia que se aplican sobre la

muestra. Cada pulso produce una rotación de la magnetización neta de la muestra hacia una determinada dirección.

El acoplamiento escalar entre espines nucleares es el resultado de la interacción indirecta de dos núcleos a través de los electrones que forman parte del enlace entre ellos. Se produce mediante la sucesiva polarización de los espines electrónicos localizados entre dos núcleos magnetizados (Figura 1.2.2). El espectro COSY es capaz de brindar información sobre la conectividad, la distancia de enlace y los ángulos interatómicos de una molécula. Por eso, el experimento COSY se ha mantenido desde el punto de vista conceptual y teórico, como uno de los más importantes en la RMN. En la actualidad, COSY es un procedimiento de rutina que se utiliza para la determinación de los acoplamientos nucleares de toda clase de compuestos.

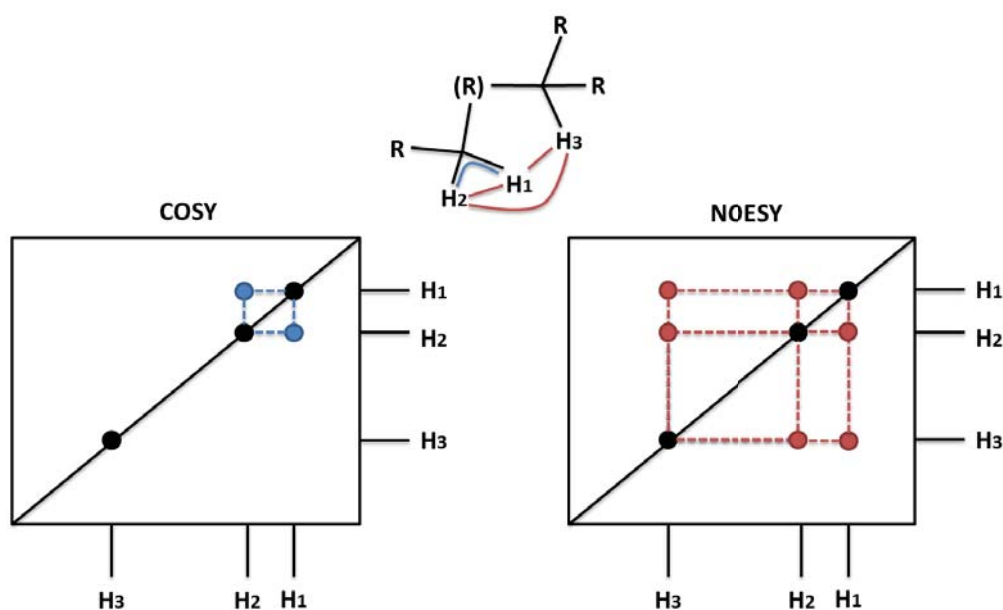


Figura 1.2.5 Representación esquemática de los espectros bidimensionales homonucleares COSY (izquierda) y NOESY (derecha). En la zona superior de la imagen está representada la estructura de la molécula hipotética a partir de la cual se obtienen los dos espectros. En azul se muestra el acoplamiento escalar y en rojo las interacciones NOE.

Además de la espectroscopia de correlación, existen otros tipos de espectroscopias bidimensionales. Una de ellas es la espectroscopia bidimensional basada en el efecto Overhauser nuclear (NOESY). La diferencia de esta espectroscopia respecto a COSY radica en las interacciones que se observan; las cuales, en este caso son producto de la relajación por el efecto NOE. El espectro resultante es similar a COSY, en ambas se encuentran picos diagonales y picos de cruce (Figura 1.2.5); pero la diferencia radica en los picos de cruce, que en el NOESY son producto del acoplamiento entre núcleos espacialmente cercanos. La información espacial que suministran los experimentos bidimensionales basados en el efecto NOE las convierte en herramientas muy útiles para el estudio de biomoléculas (Bax and Grzesiek 1993).

1.2.3 Efecto Nuclear Overhauser transferido

Es bien conocido que el efecto NOE es extremadamente útil en la determinación de la estructura 3D de las moléculas en solución (Neuhaus and Williamson 1989). En particular, ha sido muy ventajosa su aplicación en el análisis de la estructura de las proteínas (Bax 1989, Bax and Grzesiek 1993). Al tener una molécula pequeña unida a un receptor, debido al intercambio entre el estado unido y libre se empiezan a observar otra clase de NOEs (Tr-NOE). Los cambios que se observan en los NOEs debidos a la transferencia de saturación entre dos estados son las bases para una variedad de esquemas experimentales que están dirigidos a detectar y caracterizar la unión de complejos biológicos.

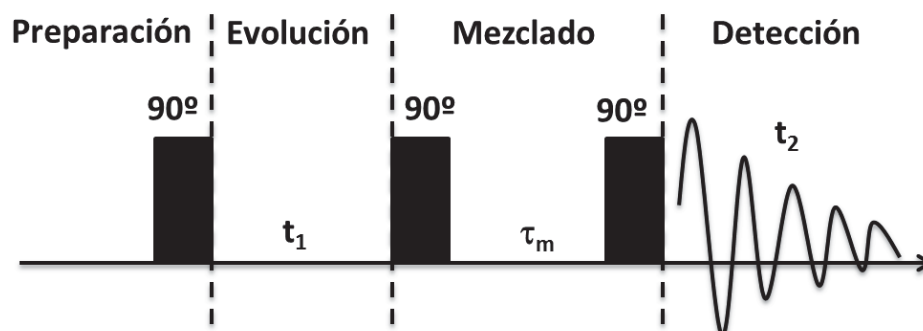


Figura 1.2.6 Esquema básico de la espectroscopia 2D de intercambio descrita en el trabajo de J. Jeener (Jeener, Meier et al. 1979).

Por su importancia, comentaremos brevemente el primer estudio relacionado con la transferencia de saturación entre dos estados desarrollado en 1979 por J. Jeener (Jeener, Meier et al. 1979). En este estudio se propuso un esquema de saturación subdividido en cuatro etapas (Figura 1.2.6). En la primera etapa se crea una magnetización transversal a través de un pulso de radiofrecuencia de 90° . Un pulso de 90° produce una rotación neta de los espines nucleares hacia el plano transversal del campo magnético B_0 . En la etapa siguiente, al vector de magnetización se le permite rotar a su frecuencia de resonancia característica por el período de evolución t_1 . Seguidamente, al inicio de la etapa de mezclado, se impone un nuevo pulso de 90° . El proceso de transferencia de magnetización vía efecto NOE tiene lugar durante esta etapa. Finalmente, se aplica sobre la muestra, un tercer pulso de 90° para rotar la magnetización longitudinal hacia el plano xy y así poderla medir.

El experimento de Tr-NOE (TR-NOESY) permite determinar la conformación bioactiva de moléculas pequeñas en un intercambio rápido con una macromolécula y emplea una secuencia de pulsos estándar básicamente similar a la descrita por J. Jeener (Figura 1.2.6). El efecto Tr-NOE fue descrito originalmente por Bothner-By en 1972 (Balaram, Bothner-By et al. 1972, Balaram, Bothner-By et al. 1972) y es el resultado de acoplar dipolarmente

espines vecinos a través de un intercambio químico con una macromolécula (Figura 1.2.7 Ilustración de la variación del efecto NOE en función del tiempo de correlación y la frecuencia del espectrómetro.). Debido a la información que se puede obtener a partir de este experimento, desde sus inicios los Tr-NOEs se consideraron muy útiles en la descripción de la estructura 3D de los complejos proteína-ligando (Poveda and Jimenez-Barbero 1998, Jimenez-Barbero, Asensio et al. 1999, Krishna and Berliner 1999).

Las señales de Tr-NOEs dependen de los tiempos de correlación rotacionales de las moléculas (τ_c) en su estado libre y unido. En los casos que se tengan moléculas con un peso molecular bajo ($M_w < 1000-2000$ Da), con tiempos de correlación cortos, los efectos NOE son positivos. Para pesos moleculares intermedios (M_w entorno a 2000 Da) se pueden obtener NOEs poco intensos o simplemente no obtener señal. Sin embargo, cuando estamos en presencia de moléculas de gran tamaño encontraremos NOEs intensos y negativos (Figura 1.2.9).

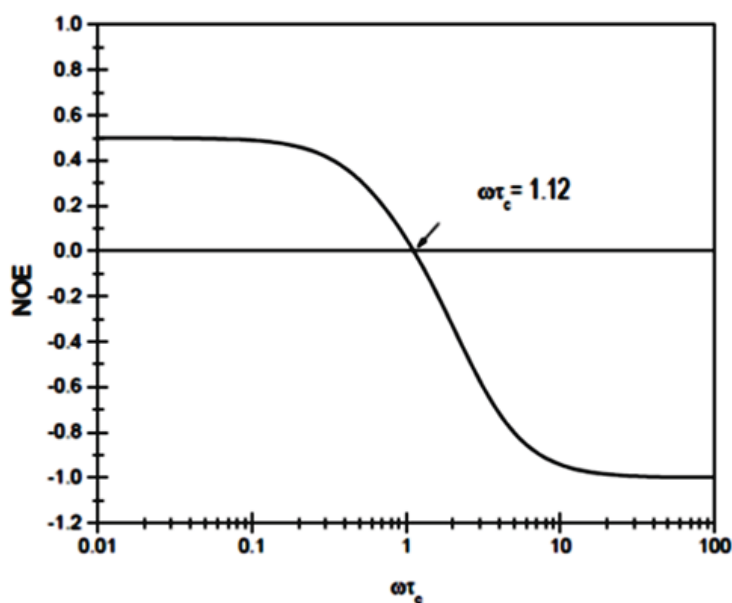


Figura 1.2.7 Ilustración de la variación del efecto NOE en función del tiempo de correlación y la frecuencia del espectrómetro.

Particularmente, en los casos donde se tienen pequeñas moléculas enlazadas a moléculas de gran tamaño y peso, la molécula ligada adopta el mismo comportamiento que exhiben los NOEs de la macromolécula al formar parte de ella. Consecuentemente, los ligandos unidos presentan fuertes NOEs negativos llamados Tr-NOEs. Los Tr-NOEs son el reflejo de la unión entre el ligando y la macromolécula, y por esto, los NOEs relacionados con la unión se pueden distinguir con facilidad en los espectros de TR-NOESY por su tamaño y signo (Figura 1.2.8) (Meyer and Peters 2003). Además, se puede discriminar entre los NOEs relacionados con las moléculas del ligando libre y los Tr-NOEs mediante el ajuste del tiempo de mezclado (τ_m) (Figura 1.2.6). Esto se debe a que el tiempo requerido para alcanzar la máxima intensidad de los Tr-NOEs está en el rango de 50 a 100 ms para el estado enlazado; mientras que para las moléculas libres el tiempo es de cuatro a 10 veces mayor (Meyer and Peters 2003). Por lo tanto, el efecto máximo en los Tr-NOEs se alcanza a tiempos de mezclado significativamente más cortos que en el estado libre. Esta circunstancia permite ajustar los parámetros experimentales con el fin de aislar los Tr-NOEs, debido a la formación del complejo molecular de los NOEs del ligando en solución.

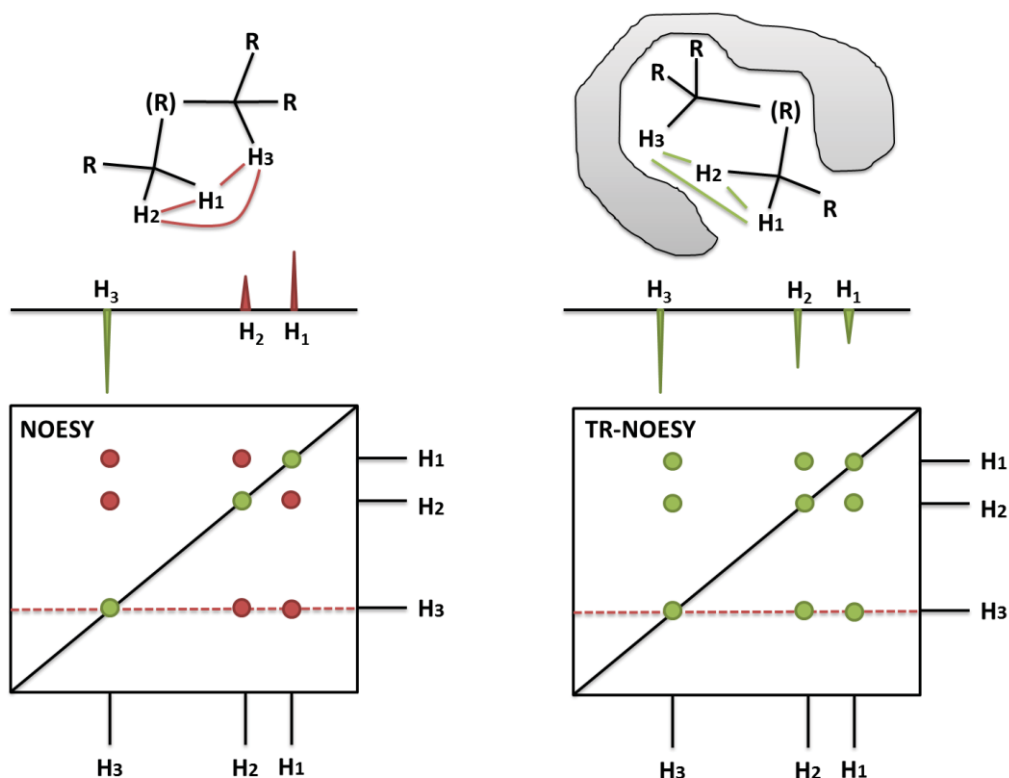


Figura 1.2.8 Izquierda, representación esquemática de un espectro NOESY. Los picos diagonales y los de cruce tienen diferentes signos. Derecha, representación esquemática de un espectro TR-NOESY. Los picos diagonales y los de cruce tienen los mismos signos. El tamaño relativo de los picos se escogió respetando la cercanía de los protones en el estado enlazado y libre. La superficie azul representa la proteína.

La dependencia de los Tr-NOE con el tiempo de correlación se explica a partir de la descripción matemática del efecto NOE. Aunque en la ecuación 1.2.4 el efecto NOE está en función de las probabilidades de transición entre estados, teniendo en cuenta ciertas consideraciones, se puede describir en función de los tiempos de correlación de las moléculas (τ_c):

$$\eta_{max} = \frac{I_i - I_i^0}{I_i} = \frac{5 + \omega_0^2 \tau_c^2 - 4\omega_0^4 \tau_c^4}{10 + 23\omega_0^2 \tau_c^2 - 4\omega_0^4 \tau_c^4} \quad 1.2.6$$

En la ecuación 1.2.6, ω_0 es la frecuencia del espectrómetro y τ_c es el tiempo de correlación rotacional de la molécula. Como no se ha comentado hasta ahora, es preciso mencionar que el tiempo de correlación rotacional (τ_c) de una molécula es el tiempo promedio que se demora dicha molécula en girar un radian alrededor de su eje. Las moléculas pequeñas tienen movimientos más rápidos y tiempos de correlación cortos; para estos casos, la ecuación 1.2.6 proporciona valores positivos. Por el contrario, en el caso de las macromoléculas, el aumento del tiempo de correlación rotacional da NOEs negativos. Pero, si las moléculas pequeñas están en constante intercambio químico con una macromolécula, como se ha comentado anteriormente, en el estado unido adquieren el tiempo de correlación de la macromolécula y sus NOEs son negativos (Figura 1.2.8, derecha). En general, no se ven las señales de las macromoléculas en los experimentos de TR-NOESY; ya que, la relación ligando:proteína que se emplea es muy alta.

Una vez llevado a cabo al análisis espectral, las intensidades de las señales de Tr-NOEs son extraídas, con el fin de evaluar sus valores y proceder a realizar el análisis conformacional de la molécula ligada. Como se ha comentado en el apartado del efecto NOE (ver apartado 1.2.1), la intensidad de la señal depende de la distancia entre protones en el estado enlazado. La dependencia entre la intensidad de los Tr-NOE y la distancia entre los espines es directamente proporcional ($1/r^6$). Por lo tanto, en un espectro bidimensional se esperan señales relativamente intensas para protones que en la estructura presentan separaciones que como media sean menores a 5 Å. En los casos en que la molécula sea muy flexible, la distancia entre los espines nucleares varía frecuentemente, provocando en consecuencia la disminución de la señal de Tr-NOE.

La información extraída de un experimento de TR-NOESY es producto de la combinación de todas las conformaciones energéticamente accesibles. Por consiguiente, la información contenida en los valores de Tr-NOEs proviene de

todas las conformaciones energéticamente posibles que adopta el ligando enlazado. Es bien conocido que las conformaciones de menor energía estarán más pobladas; y por tanto, a partir de los Tr-NOEs se puede obtener la conformación o conformaciones energéticamente más probables de la molécula en la unión con el receptor.

Los datos de Tr-NOEs pueden ser utilizados como restricciones de distancia para resolver la estructura 3D de los ligandos en su mínimo de energía conformacional. Para este fin se utilizan protocolos de modelado molecular que incluyen un paso de exploración conformacional, seguido de una optimización geométrica (Poveda and Jimenez-Barbero 1998, Bernardi, Potenza et al. 2002, Jiménez-Barbero, Canales et al. 2006, Canales, Matesanz et al. 2008).

Entre los métodos de generación conformacional se pueden encontrar los que se basan en el método de Monte Carlo (Allen and Tildesley 1989), los algoritmos genéticos (Holland 1975, Michalewicz 2013) o los modos normales de vibración molecular (Kolossváry and Guida 1996, Keserü and Kolossváry 2001). Por otra parte, existe una gran variedad de métodos de optimización geométrica; entre los más rápidos se pueden encontrar los métodos de gradiente como los de Descenso de Máxima Pendiente (Rotkiewicz and Skolnick 2008) o los de Gradiente Conjugados (Schlegel 1982, Morales and Y 2000). Pero, si se desea una optimización geométrica más rigurosa, se utilizan los métodos de segunda derivada como el de Newton-Raphson (Schlegel 2011) o el del Seguimiento del Vector Propio (Shapiro and Fan 1995). En ocasiones, con el fin de mejorar la eficiencia global de la optimización, se combinan ambas metodologías en una misma optimización geométrica.

En cualquier caso, los métodos necesitan de una metodología que describa las energías entre los átomos; metodología que puede provenir de una aproximación clásica o mecano-cuántica. En la aproximación clásica se usa un campo de fuerzas que describe la energía potencial en el estado fundamental de los átomos como función de la distancia interatómica (Allen and Tildesley 1989). En esta aproximación, se ignora el costoso cálculo de los movimientos

electrónicos y, por tanto, se puede utilizar en la descripción de sistemas con una alta complejidad estructural. Por el contrario, cuando se utiliza una aproximación mecano-cuántica, se tienen en cuenta todos los electrones, hecho que hace que los cálculos demanden mucho tiempo de cómputo y, aunque sean más rigurosos, solo se utilizan en sistemas pequeños. En ciertas ocasiones se utiliza la aproximación híbrida QM/MM, que combina la exactitud de la mecánica-cuántica con la rapidez de la mecánica molecular clásica. La notación QM/MM proviene de las siglas en ingles “*quantum mechanics/molecular mechanics*”. El método híbrido fue introducida en 1976 por Warshel y Levitt (Warshel and Levitt 1976).

La descripción de la energía a nivel clásico se realiza a través de funciones parametrizadas que se calculan a partir de estudios mecano-cuánticos o aproximaciones empíricas. Los campos de fuerzas clásicos tiene una alta especialización; pero por la importancia del estudio estructural de los complejos biológicos, existen diversos campos de fuerzas dirigidos a la descripción de proteínas (Jorgensen and Tiradorives 1988, Jorgensen, Maxwell et al. 1996, Meagher, Redman et al. 2003, Wang, Wolf et al. 2004, Bjelkmar, Larsson et al. 2010), cadenas de ácidos nucleicos (Cieplak, Cornell et al. 1995, Cornell, Cieplak et al. 1995), moléculas orgánicas (Casewit, Colwell et al. 1992, Cornell, Cieplak et al. 1995, Vanommeslaeghe, Hatcher et al. 2010) y carbohidratos (Grootenhuis and Haasnoot 1993, Hatcher, Guvench et al. 2009).

Al resolver la estructura tridimensional del ligando, es posible realizar un análisis cuantitativo de la estructura basado en los Tr-NOEs. La clave radica en lograr una formulación matemática que relacione los Tr-NOEs con las distancias entre protones y respete las condiciones experimentales. Con este fin, en 2002 se formuló la aproximación matemática de la Matriz de Relajación Completa e Intercambio Conformacional Completo (CORCEMA) para experimentos de TR-NOESY (Krishna and Moseley 2002). El CORCEMA tiene en cuenta la relajación entre toda la red de protones del complejo molecular y

el intercambio entre estados del ligando, permitiendo calcular el espectro bidimensional de TR-NOESY a partir de la estructura del complejo molecular.

El CORCEMA fue desarrollado a partir de los tratamientos matemáticos que explican el intercambio conformacional entre múltiples estados y la transferencia de la magnetización entre ellos (Krishna, Agresti et al. 1978, Krishna, Goldstein et al. 1980, Lee and Krishna 1992, Moseley, Curto et al. 1995). El método parte del modelo simplificado de dos espines aislados que describió Solomon en sus ecuaciones (ecuaciones 1.2.2, 1.2.3 y 1.2.4) (Solomon 1955) y, para poder caracterizar todos los procesos de transferencia de magnetización, construye dos matrices: la matriz de relajación completa (**R**) y la matriz de intercambio conformacional completa (**K**). En la matriz de relajación completa se describen las interacciones dipolares entre todos los espines del sistema; mientras que, en la matriz de intercambio conformacional se describen las transferencias de saturación entre los diferentes estados.

La aproximación CORCEMA aumenta el potencial del experimento de TR-NOESY. Pero, para obtener una predicción completa de la unión del ligando mediante técnicas de RMN, se necesita acoplar los datos de TR-NOESY a otro tipo de método que permita conocer con rigurosidad la orientación de la molécula en el sitio de unión. Existen varias técnicas basadas en RMN dedicadas a caracterizar la unión de los complejos moleculares (Meyer and Peters 2003, Wagstaff, Taylor et al. 2013), entre los que se puede encontrar la Diferencia de Transferida de Saturación (STD-NMR). El STD-NMR es un técnica de RMN robusta que extrae, del complejo molecular, la información estructural relacionada con la interacción entre el ligando y el receptor. Por su importancia, en el siguiente apartado se ahondará en los fundamentos de esta técnica.

1.2.4 Diferencia de Transferencia de Saturación

La Diferencia de Transferencia de Saturación o STD-NMR por su acrónimo en inglés ("*Saturation Transfer Difference*"), es un método de RMN enfocado a caracterizar la unión entre macromoléculas y compuestos pequeños con un intercambio rápido entre estados. Esta metodología goza de múltiples ventajas; entre las más atractivas se encuentran la simplicidad del procedimiento experimental, el trabajo con bajas concentraciones de proteínas, la ausencia de marcaje isotópico, la casi inexistencia de límites en el tamaño del receptor y la facilidad con que se analizan los resultados. Todas estas características hacen del STD-NMR una técnica viable y potente. En la actualidad, se utiliza tanto en la investigación académica (Vogtherr and Peters 2000, Yuan 2005, Haselhorst, Blanchard et al. 2007) como en la industria farmacéutica (Peters and Meyer 2001).

El método es muy polifacético y con él se han llegado a determinar constantes químico-físicas (Keller, Picot et al. 1979, Smith 1979, Cheshnovsky and Navon 1980, Angulo, Enríquez-Navas et al. 2010), elucidar estructuras de complejos formados por carbohidratos (Poveda, Asensio et al. 1997, Poveda and Jimenez-Barbero 1998, Jimenez-Barbero, Asensio et al. 1999, Jayalakshmi, Biet et al. 2004, Yuan 2005, Szczepina, Bleile et al. 2011), agentes anticancerígenos (Antúnez-Mojica, Rodríguez-Salarichs et al. 2016, Barra, Jiménez et al. 2016, Jantsch, Nieto et al. 2016) y antibióticos (Marcelo, Huecas et al. 2013). Desde su creación, el STD-NMR se convirtió en una técnica relevante en la caracterización de complejos moleculares de proteínas y ligandos (Cayley, Albrand et al. 1979, Birdsall, Feeney et al. 1980).

Ya a finales del siglo XX, el STD-NMR se empezó a utilizar en el cribado experimental de fármacos (Viegas, Manso et al. 2011, Wagstaff, Taylor et al. 2013, Barra, Jiménez et al. 2016). Esta nueva área fue posible por el método de Mayer y Meyer (Mayer and Meyer 1999, Meyer and Peters 2003). El método estaba dirigido al reconocimiento de la o las moléculas capaces de unirse a una

proteína de interés a partir de una librería de fármacos. Con esta nueva aplicación, el STD-NMR fue catapultado hacia la industria farmacológica (Peters and Meyer 2001).

El avance tecnológico en los siguientes años fue secundado por la reestructuración y evolución de los métodos espectroscópicos. Como técnica, su uso en la caracterización estructural de complejos moleculares es cada vez más extendido; con ella, se han logrado caracterizar sistemas moleculares que no ha sido posible resolver mediante DRX o crio-ME. Por otro parte y de carácter más novedoso, gracias a que los espectros de STD-NMR se pueden obtener en un medio biológico, se ha empezado a utilizar en el reconocimiento molecular de complejos proteína-ligando en experimentos “*in-cellulo*” (Assadi-Porter, Tonelli et al. 2008, Maggioni, von Itzstein et al. 2008, Yanamala, Dutta et al. 2010, Nietlispach and Gautier 2011, Venkitakrishnan, Benard et al. 2012, Hayafune, Berisio et al. 2014).

Anteriormente comentamos que la metodología experimental es simple y robusta, pero hasta ahora solo hemos destacado las ventajas y aplicaciones de la técnica. Como nuestro trabajo está estrechamente relacionado con el STD-NMR y su aplicación en la resolución estructural de complejos moleculares; el método se comentará con más detalle.

Un experimento de STD-NMR consta de tres pasos fundamentales (Figura 1.2.9). En el primer paso se procede a obtener un espectro monodimensional de Resonancia Magnética Nuclear protónico (^1H -RMN) del ligando en una disolución compuesta por el o los ligandos en presencia de su proteína diana. En este experimento, la frecuencia de irradiación de la secuencia de pulsos de radiofrecuencia (B_1) se fija a valores lejanos de todas las frecuencias de precesión de los protones del ligando y la proteína (~ 100 ppm), para así tener la referencia de las intensidades del espectro sin saturación. El espectro que resulta de este experimento será utilizado como referencia en el último paso del método.

En un segundo paso, se irradian selectivamente los protones de la proteína de la misma solución, imponiendo un pulso de radiofrecuencia entorno a las frecuencias de precesión de los hidrógenos metílicos o aromáticos. La frecuencia del pulso tiene que tener una diferencia de al menos 1 ppm con respecto a la señal más cercana del ligando, y usualmente, la región se escoge a partir del estudio del espectro del ligando en estado libre. Se suele utilizar en los experimentos la región del espectro entorno a -1.5 ppm si se irradian los metilos de la macromolécula. (Meyer and Peters 2003).

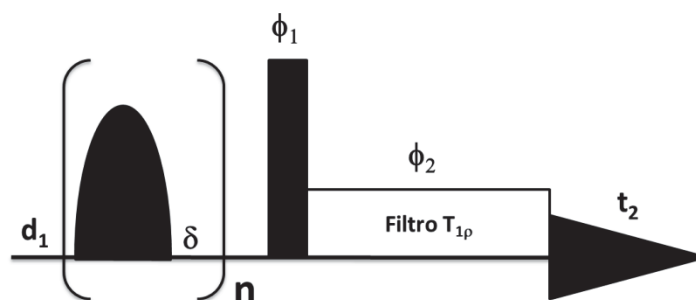


Figura 1.2.9 Secuencia de pulsos descrita por Bernd Meyer and Thomas Peters para el experimento STD-NMR (Meyer and Peters 2003). El pulso selectivo se compone de un pulso gaussiano de 50 ms y un tiempo de espera entre pulsos de $\delta=1\text{ms}$. La intensidad del pulso gaussiano de saturación selectiva es $\gamma B_1=86\text{ Hz}$. n es el número de pulsos y típicamente es igual a 40. La frecuencia del pulso selectivo es modificada al finalizar cada secuencia de pulsos entre -1,5 ppm (magnetización de los protones metílicos de la proteína) y 100 ppm. d_1 es el tiempo que se espera entre las distintas acumulaciones para que las señales relajen. ϕ_1 y ϕ_2 son las fases de los pulsos. t_2 es el tiempo de adquisición. $T_{1\rho}$ es un pulso de bloque de espín (*“spin-lock”*) de 30 ms y fase ϕ_2 que se utiliza para eliminar las señales anchas debidas a la proteína. En este procedimiento de STD-NMR, la substracción se realiza al finalizar las secuencias de pulsos.

Cuando la proteína es globular, la difusión de la magnetización se transmite homogéneamente por toda la proteína. Tras al pulso de saturación selectiva, en un corto período de tiempo, todos los protones de la proteína se encontrarán saturados y la saturación también se transferirá a los protones de las moléculas enlazadas a la macromolécula.

Por lo general, los pulsos de radiofrecuencias se aplican unos ~ 2 s (Angulo, Enríquez-Navas et al. 2010), ocurriendo en el transcurso de este tiempo múltiples intercambios químicos entre las moléculas ligadas y las macromoléculas. En cada intercambio, las moléculas ligadas se disocian y moléculas no saturadas provenientes de la solución pueden ocupar los sitios de unión libres (Figura 1.2.10). Si los compuestos tienen una disociación rápida (>1 Hz) (Meyer and Peters 2003, Wagstaff, Taylor et al. 2013), mediante el intercambio químico la magnetización se transferirá eficientemente desde la proteína hacia las moléculas del ligando en solución (Figura 1.2.10).

Las moléculas de ligando disociadas, por su pequeño tamaño, necesitan tiempo más largos para relajar totalmente su saturación. El tiempo está condicionado por el tamaño de los ligandos y se encuentra en el orden de los segundos. La combinación de los intercambios químicos y los largos tiempos de relajación induce la acumulación de la magnetización en solución. Si además consideramos relaciones proteína-ligando en un rango de 1:10 a 1:1000 (Jayalakshmi, Biet et al. 2004, Krishna and Jayalakshmi 2008), logramos amplificar el efecto acumulativo de la saturación en solución, lo que nos permite mejorar la relación señal-ruido de los espectros.

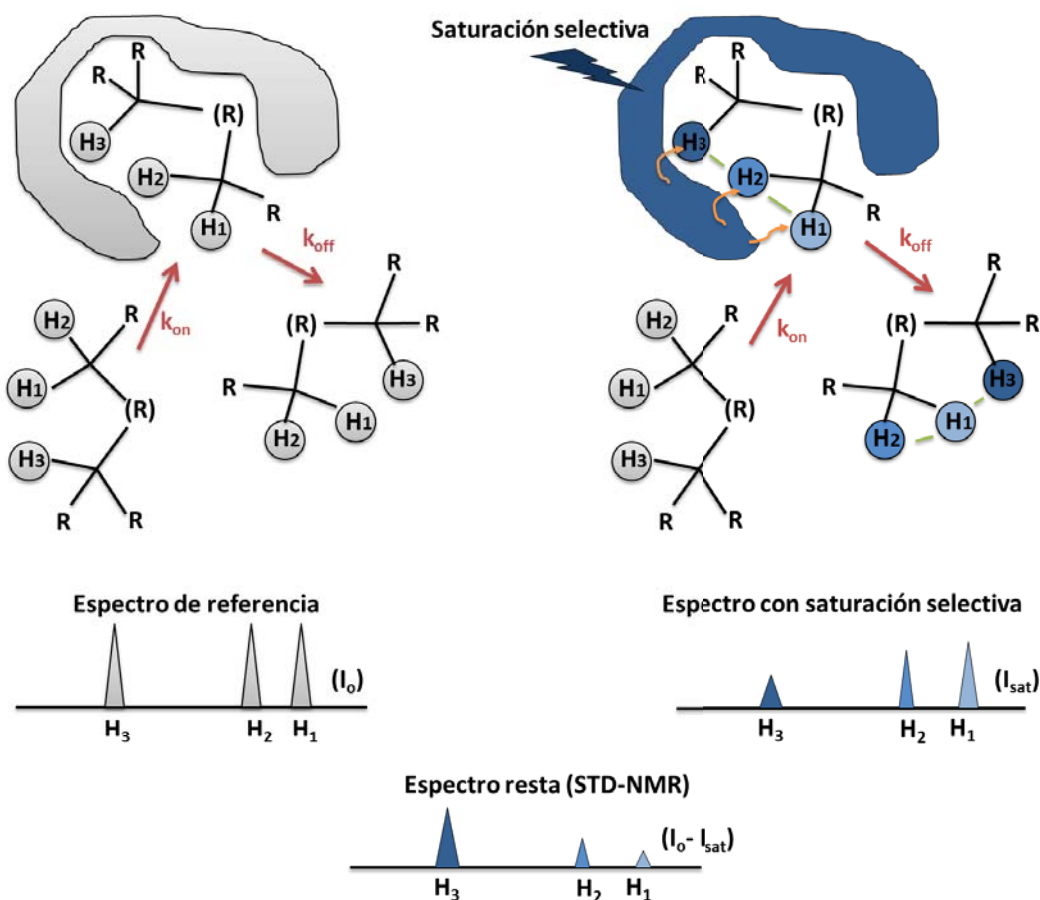


Figura 1.2.10 Representación esquemática de los diferentes pasos de un experimento STD-NMR. En la parte superior de la imagen se encuentran representadas las uniones entre la proteína y el ligando en un experimento estándar de ^1H -RMN (izquierda) y cuando se satura selectivamente los protones de la proteína (derecha). Además, se ilustran los procesos de relajación del sistema: intercambio químico (rojo), relajación intermolecular (naranja) y relajación intramolecular en el ligando (verde). La relajación intramolecular en la proteína no se representa. En la parte inferior de la figura se ilustran los espectros que se obtendrían a partir de los dos pasos. Además, se muestra el espectro diferencia (STD-NMR). La saturación protónica se representa con tonalidades de azules, siendo representados los protones más saturados con un azul intenso y los protones menos saturados con un azul claro. En ausencia de saturación los protones se representan con un morado claro.

En el tercer paso se procede a restar los espectros del primer y segundo paso para obtener el espectro STD-NMR (Figura 1.2.10). Al depender los valores de STD-NMR de la distancia entre los protones del ligando y la superficie de la proteína, se pueden caracterizar a partir de sus intensidades la estructura de los complejos moleculares. Hasta el momento no existe ninguna aproximación matemática que nos permita resolver directamente la estructura 3D del complejo molecular a partir de los datos de STD-NMR. En su ausencia, varias metodologías de modelado molecular son utilizadas para predecir la orientación de la estructura. En un procedimiento estándar, se utilizan algoritmos de acoplamiento molecular (docking) combinados con algoritmos de exploración conformacional y optimización estructural (Canales, Rodríguez-Salarichs et al. 2011, Matesanz, Rodríguez-Salarichs et al. 2011, Jantsch, Nieto et al. 2016). La modelación molecular permite explorar el espacio de coordenadas en busca de la estructura que mejor representa los datos de STD-NMR experimentales. Pero, para comparar la estructura y las intensidades de un complejo molecular, es necesario un método matemático que relacione, mediante una ecuación analítica, las coordenadas de los complejos moleculares y las intensidades de STD-NMR.

Con tal finalidad, en el 2002 la aproximación matricial CORCEMA se adaptó al experimento STD-NMR (CORCEMA-ST) (Jayalakshmi and Krishna 2002). Con la aparición del CORCEMA-ST el estudio estructural del acoplamiento molecular mejoró significativamente, logrando incluso la descripción de múltiples modos de unión entre los complejos moleculares (Angulo, Díaz et al. 2008).

La formulación del CORCEMA-ST se expresa, asumiendo un tiempo infinito entre secuencias de pulso, como

$$I(t) = I_o + [1 - \exp\{-Dt\}]D^{-1}Q \quad 1.2.7$$

donde $I(t)$ es una matriz columna que contiene la magnetización en el estado unido y libre de los protones del ligando y los protones de la proteína que no experimentan directamente el pulso de radiofrecuencia, I_0 es la matriz de magnetización correspondiente al estado de equilibrio térmico, Q es una matriz columna que contiene los términos de relajación cruzada entre los protones que experimentan directamente el pulso de radiofrecuencia y el resto de protones de la proteína y el ligando; tanto en el estado unido, como en el libre, τ es el período de tiempo que se irradian los protones de la proteína, D (matriz dinámica) es una matriz cuadrada y es igual a la suma de la matriz de velocidad de relajación completa (R) y la matriz de intercambio conformacional completa (K). A su vez, R está compuesta por las matrices de relajación completa de todos los protones en el estado libre (R_f) y unido (R_b).

$$R = \begin{bmatrix} R_f & 0 \\ 0 & R_b \end{bmatrix} = \begin{bmatrix} R_L & 0 & 0 & 0 \\ 0 & R_P & 0 & 0 \\ 0 & 0 & R_{L'} & R_{C'} \\ 0 & 0 & R_{C'}^T & R_{P'} \end{bmatrix} \quad 1.2.8$$

En la matriz, R_L y R_P son las matrices de velocidad de relajación del ligando y la proteína respectivamente en el estado libre, mientras que $R_{L'}$ y $R_{P'}$ son las matrices de velocidad de relajación del ligando y la proteína en el complejo molecular. Además, podemos encontrar la matriz $R_{C'}$ y su traspuesta $R_{C'}^T$; matrices que están relacionadas con la relajación dipolar cruzada intermolecular en el complejo. A su vez, la matriz de intercambio conformacional completo o matriz cinética (K) está compuesta por dos submatrices K_1 y K_{-1}

$$K = \begin{bmatrix} K_1 & -K_{-1} \\ -K_1 & K_{-1} \end{bmatrix} = \begin{bmatrix} k_{1L} & 0 & -k_{-1L} & 0 \\ 0 & k_{1P} & 0 & -k_{-1P} \\ -k_{1L} & 0 & k_{-1L} & 0 \\ 0 & -k_{1P} & 0 & k_{-1P} \end{bmatrix} \quad 1.2.9$$

siendo $\mathbf{k}_{1L} = k_{1L}\mathbf{1}_L$, $\mathbf{k}_{1P} = k_{1P}\mathbf{1}_P$, $\mathbf{k}_{-1L} = k_{-1}\mathbf{1}_L$, y $\mathbf{k}_{-1P} = k_{-1}\mathbf{1}_P$, donde $\mathbf{1}_N$ es la matriz unidad con una dimensión de N protones, y k_1 y k_{-1} son las velocidades cinéticas de asociación y disociación respectivamente.

La matriz CORCEMA-ST sólo permite calcular cuantitativamente los datos de STD-NMR, así que si se desea llegar a obtener una estructura es necesario utilizar métodos de modelado molecular. Pero, las metodologías de modelado molecular tienen sus propias funciones de evaluación; por consiguiente, la evaluación estructural basada en datos de STD-NMR es pasiva, ya que la evaluación se realiza al finalizar los cálculos de modelación.

En la RMN existe el factor R-NOE o factor R_e (ecuación 1.2.10). Este factor R_e es una función de error residual proveniente de los rayos-X que se utiliza en los experimentos de STD-NMR para describir la diferencia existente entre los datos experimentales y los calculados. La primer versión del factor R-NOE se utilizó en el estudio de complejos metálicos (Willcott, Lenkinski et al. 1972) y su estructura no ha cambiado mucho hasta la fecha (ecuación 1.2.10).

$$Re = \sqrt{\frac{\sum_{k=1}^{N_s} W_k (S_{exp,k} - S_{cal,k})^2}{\sum_{k=1}^{N_s} W_k (S_{exp,k})^2}} \quad 1.2.10$$

En la ecuación, $S_{exp,k}$ y $S_{cal,k}$ son las intensidades experimentales y calculadas del STD-NMR del protón k, W_k es un factor de peso igual a $1/S_{exp,k}$, y N_s es el número de protones del ligando con una adecuada señal de STD-NMR experimental. La relación inversa entre el factor de peso (W_k) es adicionada para que la función sea igualmente sensible tanto a los pequeños cambios de intensidades que sufren los protones lejanos a la superficie de la proteína como

a las desviaciones de los protones que están muy cerca de la superficie de la proteína.

El factor R_e se ha ido utilizando en la evaluación de la estructura de complejos moleculares obtenidas por modelado molecular desde que se creó el CORCEMA-ST (Jayalakshmi and Krishna 2002). Pero, no es hasta el 2004, que el factor R-NOE se aplica, por primera vez, en la determinación activa de la estructura de un complejo molecular. En el trabajo Jayalakshmi se crea un protocolo híbrido de modelado molecular que implementa el factor R_e como función de evaluación estructural en un procedimiento de exploración conformacional dedicado a optimizar la estructura de Uridina difosfato galactosa (UDP-galactosa) en el sitio de unión de galactosiltransferasa (Jayalakshmi, Biet et al. 2004). En ese trabajo se demostró que, mediante la utilización activa de los datos de STD-NMR y en particular el factor R-NOE, se puede obtener la estructura-3D de un complejo molecular guiada por datos de RMN.

La ausencia de controles experimentales que guíen los procedimientos de modelado molecular hacia la creación de la estructura-3D convierte los procedimientos mixtos en procedimientos poco fiables, ya que éstos no presentan los suficientes mecanismos de control para conocer si la estructura generada es la que mejor representa la información descrita por los datos experimentales. Una solución viable que se le puede dar a este problema sería el diseño de un algoritmo autoconsistente basado en los datos de RMN capaz de crear y evaluar activamente la estructura de un complejo molecular.

Sin embargo, la dependencia de las intensidades de STD-NMR de las condiciones experimentales, tales como las velocidades cinéticas de asociación y disociación del ligando o los tiempos de correlación de las moléculas, aumenta la complejidad del problema. Asimismo, no siempre es posible determinar experimentalmente los parámetros químico-físicos. De esta manera se hace necesaria una metodología que sea capaz de encontrar la estructura correcta sin conocer los valores de los parámetros químico-físicos.

Anteriormente se comentó que los datos de STD-NMR dependen de los acoplamientos dipolares intramoleculares e intermoleculares entre la proteína y el ligando. Aun así, los datos que se extraen de los espectros de STD-NMR no son suficientes para poder describir todas las coordenadas atómicas del complejo molecular. Por consiguiente, se hace necesaria la aplicación de otras metodologías para resolver todo el sistema molecular. Así que, en el diseño de un nuevo algoritmo basado en el experimento de STD-NMR, se tiene que contemplar la complementariedad del algoritmo con técnicas de RMN, DRX y crio-ME.

Además, los experimentos de STD-NMR pueden presentar artefactos experimentales que no se incluyen en la aproximación matricial del CORCEMA. Entre los más comunes se encuentran los procesos de relajación que no son producto del efecto NOE, la reasociación de moléculas de ligando anteriormente saturadas, múltiples poses de unión, etcétera. Todos ellos pueden provocar una divergencia en los resultados si no se tienen en cuenta. Por este motivo, el nuevo algoritmo tiene que tener en cuenta todos estos factores. Por último, la automatización del algoritmo sería bien acogida, ya que permitiría que el procedimiento estuviera al alcance del personal científico menos especializado y, por lo tanto, tuviera una mejor acogida en la comunidad científica. La combinación de todas las necesidades que tiene que cubrir el algoritmo convierte su diseño en un reto.

Aunque el STD-NMR es muy usado en la caracterización de los modos de unión de los complejos moleculares, la necesidad de un algoritmo que procure que los resultados sean cuantitativos se hace patente. Por eso, el trabajo va a versar sobre el diseño e implementación de una metodología híbrida, autoconsistente y analítica, que permita resolver la estructura-3D de complejos moleculares basada en datos de STD-NMR, Tr-NOE y datos cristalográficos. La metodología se centrará en el acoplamiento y optimización de la orientación del ligando en el sitio de unión de la macromolécula, extrayéndose las

coordenadas del ligando y la proteína a partir de experimentos de TR-NOESY y DRX respectivamente.

El trabajo estará encaminado a la creación de un programa automático y amigable para el usuario que facilite el trabajo tanto con los datos de STD-NMR como con la estructura de la proteína, el ligando y el complejo molecular. Además, para que el programa sea útil, la metodología que se implemente tiene que asegurar que la estructura real se encuentre el 100% de las veces, y además debe facilitar algún parámetro relacionado con la calidad de los resultados. Con este fin, diversas funciones de puntuación serán diseñadas e implementadas junto con el factor R-NOE en el programa automático. Además, la función error se derivará a partir de la función de dos de los factores R para obtener otro indicativo de calidad.

Posterior al diseño e implementación de las metodologías, se procederá a resolver tanto la estructura de cinco sistemas con estructuras conocidas como la estructura de un complejo molecular en un caso real de importancia para la biomedicina. En este último punto, a partir de la resolución de la estructura del complejo molecular, se llegará a caracterizar el mecanismo de asociación covalente de una familia de fármaco con propiedades anticancerígenas (Field, Pera et al. 2012) que evitan, mediante la unión covalente, los mecanismos más comunes de resistencia a fármacos (Shabbits, Krishna et al. 2001).

1.3 Caracterización de la unión colvalente entre la tubulina y la dactilolida

En el transcurso de nuestro trabajo se utilizarán repetidas veces sistemas moleculares que contienen α,β -tubulina como estructuras de referencias. Además, se resolverá su estructura en la unión covalentemente con la dactilolida, compuesto que pertenece a una familia de fármacos con propiedades anticancerígenas que evitan los mecanismos de resistencia celular. Por esto, creemos oportuno introducir la proteína brevemente.

El término tubulina se refiere a una familia de proteínas globulares altamente conservadas entre todas las células eucariotas y sus respectivos homólogos estructurales en bacterias (Jenkins, Samudrala et al. 2002, Schlieper, Oliva et al. 2005). Todas las proteínas pertenecientes a esta familia presentan un centro receptor de guanosina-5'-trifosfato (GTP) y se caracterizan por ser guanosina trifosfatasas (GTPasas). La familia de las tubulinas está formada por alfa (α), beta (β) y gamma (γ) tubulina. A su vez, las tubulinas α - y β - tienen un porcentaje de similitud del 50% y presentan un peso molecular de 55 kDa.

Los microtúbulos, por su parte, están formados por una serie de repeticiones de heterodímeros de α,β -tubulina y son parte fundamental del citoesqueleto. Los microtúbulos están involucrados en diferentes mecanismos celulares; de ellos, la interacción con el huso mitótico durante la división celular ha sido ampliamente estudiada por su relevancia para el campo de la biomedicina. La función del huso mitótico es la de anclar los microtúbulos a una estructura proteica situada sobre los cromosomas (cinetocoro). Mediante tal construcción, la célula es capaz de desplazar los cromosomas hacia los polos. Así, cuando la célula se divide, logra un reparto equitativo de los cromosomas. A causa de la función de la tubulina en la división celular, esta proteína constituye una de las dianas más comunes en terapias químicas contra el cáncer (Bhalla 2003, Honore, Kamath et al. 2003, Mani, Macapinlac et al. 2004).

Los compuestos químicos cuya diana son los microtúbulos celulares y producen la parada del ciclo celular en mitosis se les denomina agentes antimitóticos. El modo de unión de éstos puede variar, así como el efecto producido sobre los microtúbulos; ya que tanto el heterodímero de α,β -tubulina como sus formas poliméricas presentan diversos sitios de unión a fármacos. La ocupación de cada uno de los sitios por ligandos puede causar modificaciones estructurales que repercuten en la dinámica de los microtúbulos celulares.

Algunos fármacos estabilizan los microtúbulos, otros favorecen la forma dimérica y por último, existe un grupo que tiende a polimerizar la tubulina en agregados que impiden la formación de los microtúbulos (Checchi, Nettles et al.

2003). Al final, según el efecto que produzcan, se clasifican como agentes desestabilizadores de microtúbulos (MDA) o en agentes estabilizadores de microtúbulos (MSA). Los fármacos de ambos grupos, al administrarse a las concentraciones adecuadas, son capaces de suprimir la dinámica microtubular, bloquear la progresión de la mitosis e inducir la apoptosis (Jordan 2002, Jordan and Wilson 2004).

En particular, los agentes estabilizadores de microtúbulos son compuestos que al unirse a los microtúbulos favorecen su crecimiento y estabilización (Checchi, Nettles et al. 2003). La variación en la estabilidad de los microtúbulos altera, como se comentó anteriormente, la dinámica en la célula, que a su vez produce un bloqueo en el ciclo celular e induce la apoptosis (Jordan, Wendell et al. 1996). Por su importancia, los mecanismos de acción de estos compuestos están en constante estudio, lo que permite la obtención de nuevas drogas con propiedades antimitóticas mejoradas.

La dactilolida sintética es el enantiómero del producto natural extraído de la esponja "*Dactylospongia sp.*" (Cutignano, Bruno et al. 2001). Ambos compuestos son macrólidos con estructuras absolutamente opuestas. Tienen un estereocentro localizado en el átomo de carbono enlazado al grupo aldehído y un anillo 4-metilen-tetrahidropirano fusionado a un núcleo macrocíclico de lactona (Figura 1.3.1). A partir de ahora, nos referiremos a la dactilolida sintética solamente como dactilolida (DAC).

La dactilolida es un agente estabilizante de microtúbulos que se une covalentemente a la β -tubulina, y cuando se combina con células cancerígenas exhibe cierta citotoxicidad (Cutignano, Bruno et al. 2001). La unión covalente de fármacos anticancerígenos permite la evasión por parte de éstos de algunos de los mecanismos más comunes de resistencia tumoral (Kavallaris 2010). La resistencia a fármaco es una de las causas fundamentales de muerte en pacientes con cáncer (Siegel, Miller et al. 2015, Torre, Bray et al. 2015), así que la incorporación de compuestos como la dactilolida en terapias convencionales

puede reducir, en los pacientes, los efectos causados por la resistencia a fármacos (Buey, Calvo et al. 2007, Yang, Barasoain et al. 2007).

Por su importancia, las propiedades de la dactilolida como agente antitumoral fueron estudiadas recientemente por nuestro grupo mediante ensayos bioquímicos y celulares (Field, Pera et al. 2012). El análisis de los resultados reveló que el compuesto logra modificar la estructura microtubular mediante la formación de manojos de microtúbulos largos y poco abundantes en las células de carcinoma de pulmón A549 (Figura 1.3.1-B). En las mismas micrografías, se encontraron husos mitóticos aberrantes y células micronucleadas, que en conjunto frenaron el ciclo celular en la fase G₂M y arrastraron a las células hacia la apoptosis (Field, Pera et al. 2012).

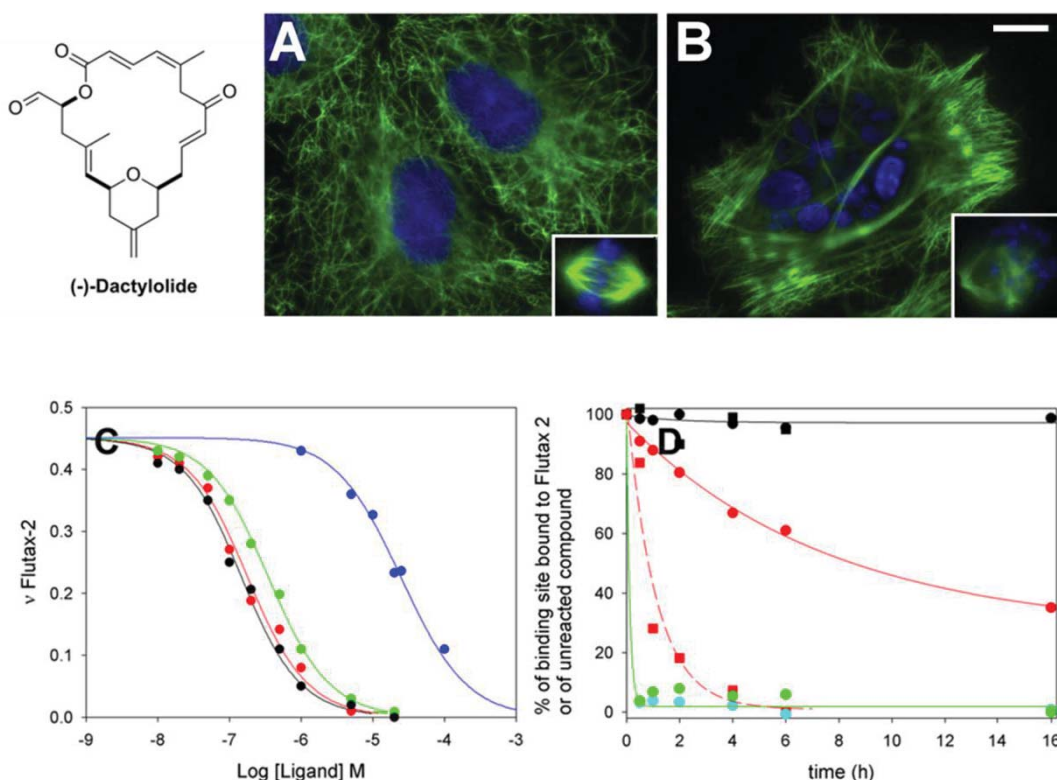


Figura 1.3.1. Estructura de la dactilolida. (A-B) Micrografías del efecto celular de la dactilolida sobre la red de microtúbulos y la morfología nuclear en el ciclo celular de la célula de carcinoma de pulmón A549. Micrografías de células

A549 incubadas por 24h con DMSO (A) y 10 μ M de dactilolida (B). Los microtúbulos (verde) se muestran mediante la inmunotinción con el anticuerpo monoclonal de α -tubulina y el DNA (azul) se muestra mediante su tinción con Hoechst 33342. Insertados, el huso mitótico en las mismas preparaciones. La barra de escala representa 10 μ m. (C) Representación de un ensayo de competición entre el Flutax-2 y MSAs por la unión por los microtúbulos. Desplazamiento de flutax-2 (50 μ M) por zampanolida (negra), docetaxel (rojo), epotilona-A (verde) y dactilolida (azul). (D) Cinética de reacción de diferentes ligandos con los microtúbulos. Círculos, Flutax-2 remanente en el sitio de unión de Taxol cuando se incubaba con DMSO (negro), ZMP natural (azul claro), ZMP sintética (verde) y DAC (rojo). Cuadrados, DAC sin reaccionar presente en solución en presencia de tampón (negro) y DMSO (rojo).

Con el fin de conocer el comportamiento molecular de la dactilolida sobre la α,β -tubulina, se realizaron diversos análisis bioquímicos bajo condiciones no polimerizantes. Primero, se analizó la habilidad de la DAC para desplazar Flutax-2 del sitio de unión. El Flutax-2 es un derivado fluorescente del paclitaxel que se caracteriza por un eficiente grupo fluoróforo y una reducida constante de asociación. Mediante este estudio, se puede conocer si la dactilolida se une a alguno de los sitios del Taxol® (sitio externo localizado en el poro tipo I o en el sitio interno en la forma microtubular (Díaz, Barasoain et al. 2003)).

El Flutax-2, la dactilolida y el heterodímero de α,β -tubulina se disolvieron bajo condiciones polimerizantes en diversas soluciones con concentraciones crecientes de dactilolida. La fluorescencia fue medida para cada solución y con las intensidades, se construyó la curva de desplazamiento de Flutax-2 (Figura 1.3.1-C). En el ensayo, las curvas del desplazamiento de conocidos agentes estabilizantes de microtúbulos fueron obtenidas para tenerlas como referencias. El desplazamiento de Flutax-2 por la dactilolida se evidenció a partir de 10⁻⁶ M de fármaco. En cambio, compuestos conocidos como el

docetaxel o la epotilona-A llegan a desplazar el Flutax-2 cuando sus concentraciones alcanzaban valores de 10^{-8} M.

Por otra parte, se obtuvieron las constantes de asociación aparente de la dactilolida a diferentes temperaturas ($K_a^{ap}(26\text{ }^{\circ}\text{C})=0,6 \pm 0,1 \times 10^6\text{ M}^{-1}$ y $K_a^{ap}(35\text{ }^{\circ}\text{C})=1,35 \pm 0,01 \times 10^6\text{ M}^{-1}$). Como se puede observar, al aumentar la temperatura aumenta la constante de asociación. Esta dependencia es opuesta a la esperada en un compuesto que forma una unión no covalente (Buey, Barasoain et al. 2005, Matesanz, Barasoain et al. 2008) y sin embargo, se asemeja al comportamiento obtenido en la unión covalente de la cicloestreptina (Edler, Buey et al. 2005). Ambos resultados parecen ser incompatibles; según las constantes de asociación, la dactilolida se comporta igual que los compuestos que se unen covalentemente a la β -tubulina y sin embargo, desplaza menos Flutax-2 que fármacos con uniones no covalentes como el docetaxel o la epotilona-A.

A causa de tal incongruencia, el análisis de la cinética del proceso de unión puede ayudarnos a conocer qué ocurre en dicha unión. Por lo tanto, se estudió la cinética de asociación mediante un ensayo de HPLC y otro de fluorescencia (Figura 1.3.1-D). Con el ensayo de HPLC se pretendió monitorizar la cinética de asociación de la unión no covalente mediante la cuantificación de la dactilolida disuelta en solución y sin reaccionar. Por otra parte, la asociación covalente fue estudiada mediante un ensayo de fluorescencia, en el cual se monitoriza, mediante la posterior adición de Flutax-2, la reducción de los sitios de unión disponibles en β -tubulina, mediante la preparación de varias soluciones en presencia de concentraciones creciente de dactilolida. A partir de estos experimentos se determinó si la unión covalente es inmediata o existe un intercambio químico rápido cuando el compuesto está asociado, pero no unido covalentemente.

A partir de este estudio se desprendió que la reacción covalente de la DAC es lenta, presentando una constante cinética aparente del orden de $0,12\text{ hr}^{-1}$ (medición realizada a través del intercambio con Flutax-2) (ver datos

representados como círculos rojos en la Figura 1.3.1-D); mientras que la reacción no covalente medida por HPLC fue 10 veces más rápida, con una constante de velocidad cinética de asociación aparente de 1 hr^{-1} (ver datos representados como cuadrados negros en la Figura 1.3.1-D). Este resultado mostró el retraso significativo que presenta el proceso de unión covalente comparado con el proceso anterior de asociación no covalente de la DAC. La diferencia entre ambas velocidades sugiere, claramente, que el compuesto asociado está bajo un intercambio rápido con el medio, reduciendo así la velocidad cinética del enlace covalente.

En el mismo estudio se caracterizó paralelamente la unión de la zampanolida (ZMP) en el sitio del Taxol®. La zampanolida es un compuesto natural altamente insaturado proveniente de la esponja marina "*Fasciospongia rimosa*" (Tanaka and Higa 1996). Su estructura está compuesta por un anillo macrocíclico de lactona de 20 miembros y una cadena lateral N-acil hemiaminal. Por tanto, los anillos macrocíclicos de la ZMP y DAC son idénticos, así que sus estructuras difieren sólo en la región de la cadena lateral (la DAC no tiene cadena lateral). Al estudiar la reacción en ambos compuestos, se encontró que la presencia de la cadena lateral favorece la unión covalente de la ZMP, ya que finaliza su reacción en solo ~30 min (ver datos representados con círculos azules claros (ZMP natural) y verdes (ZMP sintética) en la Figura 1.3.1-D) (Field, Pera et al. 2012). Así que la reacción covalente es sensiblemente más rápida en la ZMP que en la DAC ($\frac{1}{0,12} \text{ h}^{-1} = 8,33 \text{ horas}$), y esto se debe a la interacción de la cadena lateral de la ZMP con el lazo-M de la β -tubulina (Figura 5.3.2-C) (Prota, Bargsten et al. 2013). Esta interacción induce la estabilización del sitio interno de la β -tubulina, lo que provoca que la ZMP tenga una constante de asociación aparente elevada ($K_a^{ap}(26^\circ\text{C}) = 1,4 \pm 0,3 \times 10^8 \text{ M}^{-1}$).

Los procesos de estabilización molecular de la ZMP en el sitio del Taxol® permitieron su cristalización junto con el complejo T₂R (Prota, Bargsten et al.

2013). Como se comentó anteriormente, el complejo T_2R está compuesto por dos heteródímeros de α,β -tubulina unidos mediante la proteína RB3 y la enzima tubulina-tirosina ligasa (TTL). RB3 es una proteína neural que comparte el dominio de unión a microtúbulos con la estatmina Op18 (Honnappa, Cutting et al. 2003). Este se une a los dos dímeros alineados y polarizados por la cara opuesta a la del lazo-M mientras que, por su extremo N-terminal, contacta con el extremo negativo del tetrámero a modo de capuchón. Esta unión constituye la estructura conocida como T_2R y no tiene la capacidad de polimerizar ni la de desensamblarse (Gigant, Wang et al. 2005) (Figura 5.3.2-A).

En la estructura cristalográfica, se encontró que la zampanolida estaba enlazada al sitio de $\beta 2$ -tubulina, no encontrándose evidencias en las densidades electrónicas que probara su presencia en $\beta 1$ -tubulina. Esta diferencia se debe a la estructura existente en el cristal; en ésta, el sitio de unión en $\beta 1$ -tubulina se encuentra menos expuesto que el sitio localizado en $\beta 2$ -tubulina, lo que reduce el acceso a $\beta 1$ -tubulina y favorece la reacción covalente en $\beta 2$ -tubulina. Al caracterizar la asociación en $\beta 2$ -tubulina, se encontró que la zampanolida se había unido covalentemente a la HIS229. Información que ratificaba los resultados obtenidos años antes mediante estudios de espectrometría de masas acoplado a HPLC (Field, Pera et al. 2012).

En el transcurso de nuestro trabajo, se llegará a resolver la estructura de la dactilolida unida al tetrámero T_2R mediante el método híbrido obtenido con el acoplamiento de los datos de DRX y RMN. A partir de la estructura se llegará a caracterizar la unión covalente entre la familia de compuestos de la dactilolida y zampanolida a la tubulina. Las conclusiones de este trabajo posiblemente ayudarán al desarrollo de nuevos fármacos con la capacidad de evitar los mecanismos de resistencias contra fármacos en las células cancerígenas.

Hipótesis y Objetivos

2 Hipótesis

Es posible la implementación de una metodología híbrida, autoconsistente y automática basada en datos de STD-NMR, TR-NOESY y DRX que permita resolver experimentalmente la estructura de los complejos moleculares proteína-ligando.

3 Objetivos

- Diseñar una o unas funciones para evaluar la diferencia entre los datos de NMR experimentales y los calculados a partir de las coordenadas de un complejo molecular.
- Diseñar una metodología autoconsistente que permita obtener la estructura 3D de un complejo molecular mediante un algoritmo guiado por datos de NMR.
- Implementar las funciones y la metodología en un programa que obtenga la estructura de los complejos moleculares automáticamente.
- Analizar las condiciones necesarias para que el nuevo programa asegure la obtención de la estructura que mejor describe los datos de STD-NMR experimentales.
- Diseñar e implementar un método híbrido autoconsistente que permita resolver todas las coordenadas de un complejo molecular proteína-ligando.
- Caracterizar la unión covalente del complejo molecular de dactilolida y el heterodímero de α,β -tubulina.

Materiales y métodos

4 Materiales y métodos

4.1 Productos químicos

El dodecil sulfato sódico (SDS), el guanosín trifosfato (GTP) y el guanosín difosfato (GDP) fueron de Sigma. El Sephadex G25 y el DEAE Sephadex A50 fueron de Amersham. El cloruro de guanidinio fue de *United States Biochemical Corporation*. El DMSO grado espectroscopía, el DMSO-d₆ grado espectroscopía, el fosfato monosódico y disódico grado analítico, el ácido etilen-diamino-tetra-acético (EDTA), el cloruro de magnesio grado analítico, el sulfato de amonio puro y para análisis, la sacarosa grado para fines bioquímicos, el cloruro de potasio grado analítico, el acetato de uranilo fueron de Merck. El glutaraldehído grado microscopía fue de Emscope. El borohidruro sódico fue de Sigma.

4.2 Tampones

Tampón PEDTAx-x' mM GTP/GDP: 10 mM de fosfato sódico (NaH₂PO₄/Na₂HPO₄), 1 mM de EDTA, 7 mM de MgCl₂ y 1 mM de GTP/GDP a pH 7,0.

4.3 Proteínas

La tubulina empleada se purificó a partir de de cerebros bovinos de edades inferiores a 12 meses procedentes del Matadero Municipal Vicente de Lucas, en Segovia. El protocolo de purificación de tubulina empleado se basa en el protocolo descrito por Weisenberg (Weisenberg, Borisy et al. 1968, Lee, Frigon

et al. 1973) con la adición de algunas modificaciones (Andreu 2007). En cada purificación, se utilizan una media de 8 o 9 cerebros para obtener un rendimiento de 300 a 600 mg de tubulina purificada. En el procedimiento, se homogeneizan los sesos usando un homogeneizador de aspas, y posteriormente se eliminan las membranas celulares por centrifugación. A continuación, se procede a separar los componentes de la solución mediante repetidos pasos de precipitaciones salinas, precipitando la tubulina en presencia de un 43% de sulfato amónico en solución. El precipitado se resuspende, y se separa la tubulina mediante una cromatografía de intercambio iónico con DEAE-Sephadex A50. La muestra resultante se precipita de nuevo con sulfato amónico (43 %) y se resuspende en un pequeño volumen para, a continuación, cargarla en una columna de cromatografía de filtración en gel Sephadex G25. Seguidamente, se adicionan 30 mM de cloruro de magnesio a cada una de las fracciones recogidas para identificar las que contienen tubulina funcional. Las soluciones que contienen tubulina funcional se combinan y se someten a una diálisis de al menos 16 horas para eliminar el exceso de cloruro de magnesio. La solución de diálisis se prepara con 1 M sacarosa para crioproteger la tubulina. La concentración de tubulina se determina espectrofotométricamente a una longitud de onda de 275 nm ($\epsilon_{275\text{nm}}=107000 \text{ M}^{-1}\text{cm}^{-1}$), diluyendo 200 veces la solución en 6 M de cloruro de guanidinio (Andreu, Gorbunoff et al. 1984). Finalmente, la solución se distribuye convenientemente en pequeñas porciones de 5, 10 y 20mg de tubulina que son almacenadas posteriormente en nitrógeno líquido.

Antes de realizar cada experimento, las alícuotas de tubulina se descongelan rápidamente y se equilibran en el tampón deseado mediante un proceso de cromatografía de filtración en gel, siendo la fase estacionaria una columna de 0.9 x 25 cm de Sephadex G25. La presencia de tubulina en las fracciones se identifica espectrofotométricamente fijando la absorbancia a 295 nm. Posteriormente, las fracciones se centrifugan conjuntamente en un rotor TLA 120.2, utilizando una centrífuga Optima TLX de Beckman, a 50000 rpm, por 10 minutos y a 4 °C. La concentración final de tubulina se determina

espectrofotométricamente diluyendo una porción 20 veces en 10 mM fosfato sódico, 0.1 % SDS, pH 7.0.

4.4 Preparación de tubulina en condiciones no polimerizantes

Con el fin de obtener diluciones homogéneas de tubulina y GDP (Díaz and Andreu 1993), la sacarosa se elimina del tampón de almacenamiento mediante una columna semiseca de centrífuga con Sephadex G25 (0,9 x 10 cm) y equilibrada en un tampón PEDTA-1 mM GDP (ver apartado 4.2). La solución de proteína resultante se incuba en hielo durante 10 minutos en presencia de 10 mM de GDP. El exceso de nucleótido se elimina mediante cromatografía de filtración en gel, con Sephadex G25, en una columna de dimensiones 0,9 x 10 cm y en presencia de un tampón de PEDTA-1 mM GDP. Las fracciones con absorbancias superiores a una unidad a 295 nm, se combinan y posteriormente se centrifugan a 50000 rpm, por 10 minutos, a 4 °C, en un rotor TLA 120.2 y una centrífuga Optima TLX de Beckman. La concentración final de tubulina se determina espectrofotométricamente en 10 mM fosfato sódico, 0,1 % SDS y pH 7,0; como se describe en el apartado 4.3.

4.5 Preparación de muestras para experimentos de RMN

Las muestras se preparan siguiendo el protocolo descrito en el apartado 4.4 y en (Canales, Rodríguez-Salarichs et al. 2011, Field, Pera et al. 2012, Pera, Barasoain et al. 2013). En resumen, en cada experimento se prepararon 500 µL de soluciones deuteradas conteniendo 300 µM de ligando y 10 µM de tubulina, con 2% de DMSO-d₆, en un tampón a pH 7,0 que contiene NaPi 10 mM, GTP 0,1 mM y D₂O.

4.6 Experimentos de RMN

Los experimentos de STD-NMR y TR-NOESY fueron realizados por la Dra. Ángeles Canales Mayordomo en el laboratorio del Dr. Jiménez-Barbero, en el Centro de Investigaciones Biológicas de Madrid (España), tal y como se describe en (Canales, Rodríguez-Salarichs et al. 2011, Field, Pera et al. 2012, Pera, Barasoain et al. 2013). Los experimentos de RMN se llevaron a cabo a 298 K en DMSO-d₆ con espectrómetros Bruker Avance de 500 MHz, 600 MHz y 700MHz. Los experimentos de STD-NMR se obtuvieron irradiando la muestra a -1 ppm y 100 ppm con un tiempo de saturación de 2 s, y concatenando pulsos gaussianos de 50 ms separados un tiempo de 1ms. El espectro de STD-NMR se obtuvo substrayendo el espectro de saturación selectiva y el espectro unidimensional protónico (RMN-H¹). Los datos resultantes se analizaron en el programa TopSpin 3.0 de Bruker. Solo se tuvieron en cuenta los picos no solapados con una SNR > 10.

Los experimentos de TR-NOESY se adquirieron con una secuencia de pulsos con la supresión de agua Watergate y un tiempo de mezclado de 200 ms. En ellos no es necesario aplicar un pulso de bloque de espín para eliminar la señal de fondo de la proteína porque debido al tamaño de la proteína su señal no se observa. Como la adquisición de los datos demora unas horas en el transcurso de todos los experimento, fueron adquiridos diversos espectros de RMN-H¹ como control para verificar la existencia de cambios moleculares debidos a procesos de degradación.

4.7 Experimentos de cristalografía y DRX

Los experimentos de cristalografía con el compuesto PM060327 y la dactilolida unida al tetrámero de tubulina (T₂R) fueron realizados por el Dr. Andrea Prota

en el laboratorio del Dr. Steinmetz, en el Paul Scherrer Institut, Villigen PSI (Suiza), tal y como se describe en (Prota, Bargsten et al. 2013).

4.8 Cálculos de modelado molecular

El estudio conformacional de los ligandos fue realizado en ausencia de la proteína utilizando el paquete Macromodel (Mohamadi, Richards et al. 1990), el campo de fuerzas OPLS2005 (Jorgensen and Tiradorives 1988, Jorgensen, Maxwell et al. 1996, Kaminski, Friesner et al. 2001) implementado en el programa Macromodel y el método de búsqueda de muestreo torsional de pseudo-Monte Carlo (MC/MM) (Chang, Guida et al. 1989). La solvatación del agua se consideró a través del modelo de solvente continuo generalizado de Born GB/SA (Still, Tempczyk et al. 1990). La búsqueda conformacional se realizó sobre toda la estructura, utilizando 20000 iteraciones con distancias de corte para el potencial de van der Waals de 8,0 Å y el potencial electrostática de 20,0 Å. Además, se impusieron fuerzas restrictivas de 50, 100 y 200 kJ/molÅ² a los pares atómicos con señales de TR-NOESY débiles, medias y fuertes. Las estructuras resultantes se optimizaron con el algoritmo de Gradiente Conjugado Polak-Ribiere (PRCG) considerando un máximo de iteraciones de 90000 o un gradiente de energía mínimo de 0.05 kJ/molÅ⁻¹.

El acoplamiento molecular entre los ligandos y las proteínas se calculó utilizando el programa AutoDock 4.2.6 (Morris, Goodsell et al. 1998). En cada cálculo, las moléculas son consideradas rígidas. Además el espacio de coordenadas a explorar va a estar definido por el contorno de una caja cúbica centrada en la cavidad de la proteína con 126 puntos de energías y con un espaciado entre puntos de 0.375 Å. Durante una simulación, se realizaron 100 carreras del Algoritmo Genético Lamarckiano con un tamaño de poblacional igual a 150, un número máximo de generaciones de 27000 y número máximo de evaluaciones de energía de 250000. Cada carrera proporcionó un único modo de unión que posteriormente se compararon y agruparon por su RMSD.

Al conocer la orientación que tiene el ligando en la cavidad de la proteína se puede proceder con la exploración torsional guiada del ligando (“*Coordinate scan*”). En la exploración torsional del ligando se especificaron los ángulos torsionales que se desean estudiar en el Macromodel y se rotaron 360° en pasos de 5°. En cada paso torsional guiado, se les impuso una fuerza de 1000 kJ/mol a los átomos que no formaron parte del ángulo diedro para reducir su movimiento y así respetar las coordenadas iniciales. Posterior a cada rotación, se realizó la optimización de la estructura para reducir los efectos estéricos. El paso de optimización estructural se realizó con el algoritmo de Gradiente Conjugado Polak-Ribiere (PRCG) considerando un máximo de iteraciones de 90000 o un gradiente de energía mínimo de 0.05 kJ/molÅ⁻¹.

4.9 Cálculos con el CORCEMA-ST

Los cálculos de STD-NMR se realizaron con el CORCEMA-ST (Jayalakshmi and Krishna 2002). En cada cálculo, se consideran solo los residuos que estaban a 10 Å del ligando. El parámetro de orden que describe los movimientos internos del metilo se igualó a $S^2=0,85$ (Dellwo and Wand 1993, Brüschweiler and Case 1994, Moseley, Lee et al. 1997) para las interacciones entre los metilos y el resto de protones; ya que, en la interacción entre protones del mismo grupo metilo, el CORCEMA-ST consideró el parámetro de orden de $S^2=0,25$. Además, se tuvo en cuenta un tiempo infinito entre secuencia de pulsos. Por último, el factor de fuga que tuvo en cuenta la velocidad de relajación no específica debida al oxígeno paramagnético disuelto se consideró como 0,3 (Moseley, Lee et al. 1997).

Resultados

5 Resultados

5.1 Diseño de funciones de coste para la comparación de los STD-NMR experimentales y los calculados

5.1.1 Diseño de alternativas al factor R_e

El factor R-NOE o factor R_e es una función de error residual usada para evaluar la calidad de las estructuras obtenida mediante datos de RMN (Gonzalez, Rullmann et al. 1991). Su construcción es análoga a la empleada en la determinación estructural por rayos-X, con la diferencia de que en los experimentos de RMN, este factor es el encargado de describir la diferencia existente entre datos experimentales y calculados. Como se comentó en capítulos anteriores, la formulación del factor R-NOE utilizada en la evaluación de datos de STD-NMR obtenidos a partir de estructuras de complejos moleculares presenta algunos inconvenientes. En un intento de solventar sus debilidades, cinco funciones fueron desarrolladas empíricamente como alternativas al factor R_e . Tres factores R-NOE (R_t , R_{tm} y R_{twm}) fueron formulados conservando la estructura matemática del factor R_e . Mientras que, la construcción del resto de las funciones se basó en la estructura del error cuadrático medio (R_{ms-i} y R_{msw-i}).

El factor R-NOE fue especialmente adaptado para evaluar datos de STD-NMR mediante la introducción del término $W_{exp,k}$. (Jayalakshmi and Krishna 2002). $W_{exp,k}$ posibilita que el factor R_e sea sensible tanto a variaciones significativas en valores de STD-NMR pequeños, como a desviaciones en valores grandes de STD-NMR (Angulo and Nieto 2011). Como efecto negativo, el término sobrestima los datos provenientes de señales de menor relación señal-ruido (SNR).

$$R_t = \sqrt{\frac{\sum_{k=1}^{N_s} W_{cal,k} (S_{exp,k} - S_{cal,k})^2}{\sum_{k=1}^{N_s} W_{cal,k} (S_{exp,k})^2}} \quad 5.1.1$$

$$R_{tm} = \sqrt{\frac{\sum_{k=1}^{N_s} W_{cal,k} (S_{exp,k} - (S_{cal,k} + \Delta M))^2}{\sum_{k=1}^{N_s} W_{cal,k} (S_{exp,k})^2}} \quad 5.1.2$$

$$R_{twm} = \sqrt{\frac{\sum_{k=1}^{N_s} W_{cal,k} (S_{exp,k} - (S_{cal,k} + \Delta M_w))^2}{\sum_{k=1}^{N_s} W_{cal,k} (S_{exp,k})^2}} \quad 5.1.3$$

Los factores R_t , R_{tm} y R_{twm} (ecuaciones 5.1.1, 5.1.2 y 5.1.3) eliminan la sobrestimación de señales con baja relación señal/ruido sin perder la sensibilidad. En las tres funciones, el parámetro $W_{exp,k}$ del factor R_e es substituido por la inversa de la saturación calculada ($W_{cal,k} = 1/S_{cal,k}$). $W_{cal,k}$ es capaz de mantener la sensibilidad, al mismo tiempo que disminuye la sobrestimación de las intensidades experimentales. En trabajos anteriores se ha hecho alusión al término $W_{cal,k}$ (Angulo, Díaz et al. 2008), pero hasta la fecha no se ha utilizado, o comparado con el $W_{exp,k}$.

El procedimiento de obtención de datos de STD-NMR teóricos a través del método CORCEMA depende de cuatro parámetros experimentales. La obtención de todos ellos conlleva una labor experimental adicional y en algún caso, la ausencia de medios para sus determinaciones. Como se comentó en el capítulo de métodos, las velocidades de asociación y disociación del ligando y los tiempos de relajación de ambas moléculas son los parámetros a determinar. Debido al coste en tiempo y esfuerzos, en un procedimiento estándar, los parámetros se resuelven mediante aproximaciones teóricas

(Garcia de la Torre, Huertas et al. 2000, Canales, Rodríguez-Salarichs et al. 2011), ajustes paramétricos basado en la minimización del factor R-NOE (Wen 2005, Canales, Nieto et al. 2014, Jantsch, Nieto et al. 2016) o mediante ajustes empíricos (Jayalakshmi and Rama Krishna 2004).

El último tratamiento se aplica, según la experiencia del autor, para fijar arbitrariamente el valor de la constante de equilibrio y la constante de velocidad de asociación. Aunque esta aproximación no es conveniente en la mayoría de los métodos de elucidación estructural, el STD-NMR permite, de manera grosera, tener en cuenta esta consideración siempre que se cumplan ciertas condiciones (Jayalakshmi and Krishna 2002).

Los autores del CORCEMA lograron caracterizar las condiciones en el estudio de dos sistemas teóricos (Jayalakshmi and Krishna 2002). En este se concluyó que bajo un régimen de intercambio molecular rápido, el STD-NMR protónico no varía para $10^4 \leq K_a \leq 10^7 \text{ M}^{-1}$. En la mayoría de los compuestos moleculares estudiados a través de experimentos de STD-NMR se cumplen estas condiciones, por lo que la aproximación es válida. En el caso de la velocidad de asociación es muy común asumir un proceso asociativo entre moléculas que es controlado por la difusión, de forma que se suele tomar una $k_{on} \approx 10^7 \text{ s}^{-1} \text{ M}^{-1}$ (Meyer and Peters 2003).

El trabajo adicional generado en la obtención de los parámetros químico-físicos incita la búsqueda de un factor R-NOE menos dependiente de estos parámetros. Como es bien sabido, la transferencia de la magnetización entre protones está influenciada por las características físico-químicas del sistema (Strehlow and Knoche 1977, Brüschweiler and Case 1994). Por este motivo, el factor R_e no es capaz de discernir entre el aporte de la disposición espacial y el aporte paramétrico en las diferencias de los datos calculados y experimentales. En el análisis realizado en el apartado 5.1.3 se profundiza sobre el comportamiento del STD-NMR y la influencia de los parámetros cinéticos y de relajación. Los conocimientos adquiridos en este apartado fueron aplicados en la renovación del factor R-NOE para reducir su dependencia paramétrica.

Los factores R_{tm} y R_{twm} tienen un término nuevo encargado de rectificar la desviación global de la saturación calculada debida a la parametrización. En ambos se substraen la diferencia media entre la saturación experimental y calculada (ΔM). Para R_{tm} , la diferencia media se consideró aritmética (ecuación 5.1.4), mientras que en R_{twm} , la diferencia de saturación ($S_{exp,k} - S_{cal,k}$) es ponderada con $W_{cal,k}$ (ΔM_w). En ΔM_w se sustituyeron los términos de la saturación experimental y calculada por la inversa de W_k y se reestructuró para obtener la ecuación 5.1.5.

$$\Delta M = \frac{\sum_{k=1}^{N_s} (S_{exp,k} - S_{cal,k})}{N_s} \quad 5.1.4$$

$$\Delta M_w = \frac{N_s \sum_{k=1}^{N_s} (W_{exp,k} - W_{cal,k})}{\sum_{exp,k}^{N_s} \sum_{cal,k}^{N_s} W_{exp,k} W_{cal,k}} \quad 5.1.5$$

Como nombramos al principio, hemos desarrollado cinco funciones como alternativas al factor R_e . Hasta ahora, nos hemos centrado en las tres que han sido formuladas conservando la misma estructura matemática, sin embargo, también debemos considerar los otros dos factores R-NOE: R_{ms-i} y R_{msw-i} (ecuaciones 5.1.6 y 5.1.7). En la formulación de los factores R_{ms-i} y R_{msw-i} se ha optado por la estructura de la desviación cuadrática media. El cambio de paradigma se debió a un factor práctico. Ambas tienen un término de penalización, resultando más cómodo el uso de la desviación cuadrática antes que el error residual. La diferencia entre ambas ecuaciones radica en el término ΔM .

Tanto la ecuación 5.1.6 como 5.1.7 incluyen el factor para la reducción de la desviación global de la saturación calculada y tienen añadido el término de penalización $N_i \Delta M^2$. Este término penaliza los ajustes paramétricos que no

minimizan la diferencia media entre la saturación experimental y calculada. De este modo, las ecuaciones presentan un doble efecto, reducen la diferencia ocasionada por el ajuste y al mismo tiempo castigan los ajustes que producen las mayores desviaciones.

$$Rms - i = \sqrt{\frac{\sum_{k=1}^{N_s} (S_{exp,k} - (S_{cal,k} + \Delta M))^2 + N_i \Delta M^2}{N_s + N_i}} \quad 5.1.6$$

$$Rmsw - i = \sqrt{\frac{\sum_{k=1}^{N_s} (S_{exp,k} - (S_{cal,k} + \Delta M_w))^2 + N_i \Delta M_w^2}{N_s + N_i}} \quad 5.1.7$$

Estas funciones tienen la ventaja de tratar el factor de penalización con flexibilidad. La variable N_i es un número natural encargado de modular el factor de penalización. El subíndice i está definido como $i = \frac{100 \times N_i}{N_s + N_i}$, siendo N_s el total de protones saturados del sistema.

Partiendo de esta relación, si N_i se iguala a N_s , el término ΔM^2 se tiene en cuenta tantas veces como protones saturados tenga el ligando. La diferencia entre las intensidades calculadas ($S_{cal,k}$) y experimentales ($S_{exp,k}$) contribuiría solo un 50%. Por otro lado, según su definición el subíndice i valdría 50; por lo que, la nomenclatura de los factores R-NOE cambia a R_{ms-50} y R_{msw-50} .

Si consideramos $i > 50$ ($N_i > N_s$), el factor R-NOE se torna más sensible a la desviación de la saturación calculada, como consecuencia de la parametrización, que a los cambios en la saturación debidos a las coordenadas atómicas del complejo molecular.

Por otro lado, debido a la complejidad que supondría la exploración de todo el rango de valores de i , y asumiendo que uno de los objetivos de este trabajo es encontrar las modificaciones matemáticas que disminuyen la sensibilidad del factor R-NOE frente a la parametrización y que no se debe de reducir la sensibilidad del factor al cambio de coordenadas espaciales, concluimos que debemos de considerar la evaluación de R_{ms-i} y R_{msw-i} solo cuando $i = 50$.

Aun así, realizamos una prueba preliminar para R_{ms-i} considerando $0 \leq i \leq 1$. Los factores R_{ms-i} y R_e fueron utilizados como funciones de puntuación en la caracterización de la unión de dos agentes estabilizantes de microtúbulos (Canales, Salarichs et al. 2011). De los resultados arrojados por las funciones, se escogieron los obtenidos por R_{ms-50} , al identificar las mismas posiciones de unión que R_e . La respuesta mostrada por R_{ms-50} , junto a la necesidad de simplificar el estudio comparativo de los seis factores R-NOE, nos reafirmó la decisión de considerar en el presente estudio $i=50$ para R_{ms} y R_{msw} .

Como resumen, tenemos el factor R-NOE tradicional y cinco derivados de nuevo diseño con la capacidad potencial de evaluar la calidad de las estructuras mediante datos de STD-NMR. Los primeros cuatro factores (R_e , R_t , R_{tm} y R_{twm}), de los seis presentados en este apartado, son sensibles tanto a variaciones significativas en valores de STD-NMR pequeños como a desviaciones en valores grandes de STD-NMR. Las últimas cuatro formulaciones del factor R-NOE (R_{tm} , R_{twm} , R_{ms} y R_{msw}) nos permiten evaluar las orientaciones de un ligando en el sitio de unión teniendo en cuenta, de manera explícita, el efecto producido por las constantes cinéticas y los tiempos de relajación. De todos ellos, los factores R_{tm} y R_{twm} son las únicas formulaciones que contienen ambas modificaciones.

5.1.2 Búsqueda de sistemas modelos para la evaluación de las funciones diseñadas

Los seis factores R-NOE van a ser estudiados en profundidad. Para poder analizarlos es necesario tener complejos moleculares cuyas estructuras atómicas y datos de STD-NMR estén resueltos a alta resolución. Adicionalmente se tiene que cumplir que las dos técnicas describan la misma estructura atómica.

En este sentido, se encontraron tres sistemas de referencia que cumplieran las condiciones establecidas. Las estructuras atómicas de los tres complejos moleculares se resolvieron por técnicas cristalográficas con una resolución menor a 2,5 Å. Dos de ellos, kifunesina (Figura 5.1.1-A) y salacinol (Figura 5.1.1-B), son pequeñas moléculas inhibidoras de glucosidasas y fueron cristalizados junto a la α -manosidasa II de la *Drosophila melanogaster* (dGMII) (Shah, Kuntz et al. 2003, Kuntz, Ghavami et al. 2005). El tercer complejo (Figura 5.1.1-C) lo forman la epotilona-A y el complejo de estatmina-tubulina (T_2R) (Prota, Bargsten et al. 2013). La epotilona-A es un agente estabilizador de microtúbulos y pertenece a una familia de drogas con potencial aplicación en la terapia contra el cáncer (Thomas, Tabernero et al. 2007). Los datos de STD-NMR de los complejos moleculares de kifunesina y salacinol fueron gentilmente cedidos por Mario Pinto y se obtuvieron en un espectrómetro Bruker AMX-600 a 295 K (Yuan 2005). Mientras que, la epotilona-A unida al heterodímero de α,β -tubulina fue caracterizada mediante datos de STD-NMR y TR-NOESY en nuestro grupo utilizando un espectrómetro Bruker AVANCE de 500 MHz a 298 K (Canales, Nieto et al. 2014).

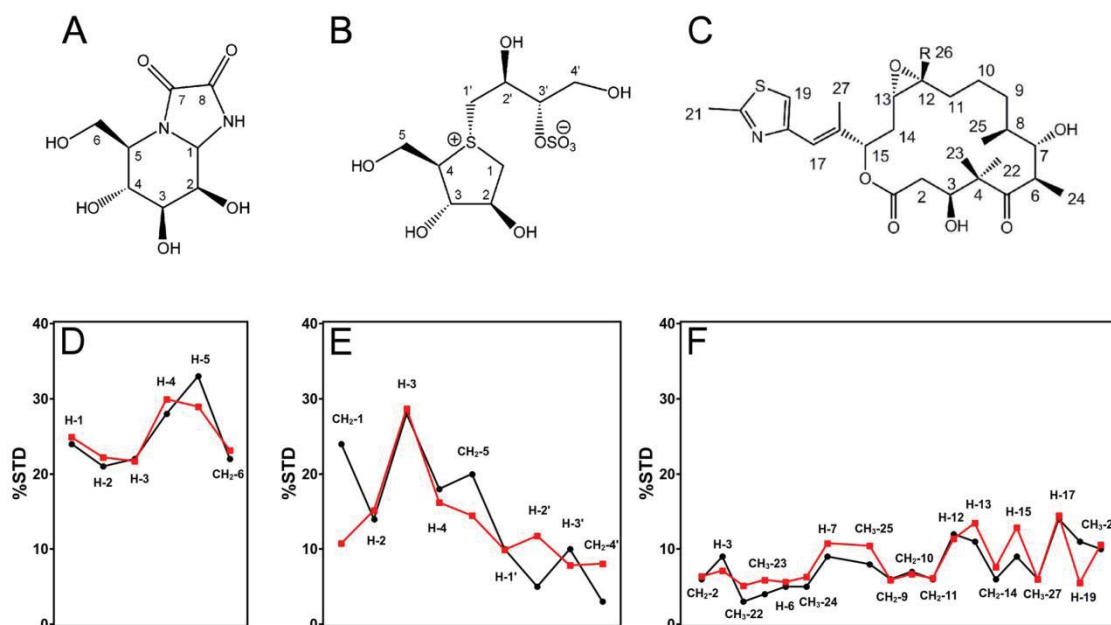


Figura 5.1.1. Estructuras químicas de los ligandos de (A) kifunesina, (B) salacinol y (C) epotilona-A. Datos de STD-NMR experimentales y calculados a partir de la estructura cristalográfica de los complejos de (D) kifunesina-dGMII (PDB: 1PS3), (E) salacinol-dGMII (PDB: 1TQS) y (F) epotilona-A- α,β -heterodímero de tubulina (PDB: 4I50). Los datos de STD-NMR experimental están representados con líneas y círculos negros. Los datos de STD-NMR calculados están representados con líneas y cuadrados rojos.

Para validar los tres sistemas antes de usarlos en nuestros análisis se han contrastado las posiciones de unión descritas por DRX con los datos de STD-NMR experimentales. Para esto, solo los amino ácidos a 10 Å de cada ligando fueron considerados en el cálculo del STD-NMR. Además, los parámetros químico-físicos fueron explorados y se escogió el grupo de parámetros que en su conjunto minimizaban el factor R_e . Finalmente, las intensidades calculadas fueron comparadas con los datos de STD-NMR experimentales (Figura 5.1.1-D, E y F). El conjunto de parámetros que minimizan la diferencia entre los datos de STD-NMR experimentales y calculados se encuentran expuestos en la Tabla 5.1.1.

Tabla 5.1.1. Conjunto de parámetros que minimizan el factor R_e en los complejos de kifunesina-dGMII, salacinol-dGMII y epotilona-A- β -tubulina.

	Kifunesina-dGMII	Salacinol-dGMII	Epotilona-A- β - tubulina
$\tau_{c-unido}$ [ns]	300	200	40
$\tau_{c-libre}$ [ns]	0.75	0.25	0.25
K_a [M ⁻¹]	10^4	10^4	10^5
k_{off} [s ⁻¹]	10^2	10^3	10

En la Figura 5.1.1 se puede observar como los datos de STD-NMR calculados describen el mismo patrón que los datos experimentales en los tres sistemas. La existencia de pequeñas desviaciones en las intensidades calculadas de algunos protones se debe a la inherente flexibilidad descrita por ciertas zonas de los ligandos. Como los cálculos se realizan a partir de un sistema de coordenadas rígido, encontrar pequeñas desviaciones en los datos de STD-NMR es común en moléculas flexibles (Canales, Nieto et al. 2014, Jantsch, Nieto et al. 2016). Por consiguiente, si nos basamos en los resultados de la comparación podemos considerar que ambas técnicas describen la misma pose de unión en los complejos de kifunesina-dGMII, salacinol-dGMII y epotilona A- β -tubulina.

5.1.3 Estudio de la variación del patrón de STD-NMR de los sistemas modelos con los parámetros químico-físicos

Para un buen análisis del STD-NMR es importante tener en cuenta su relación con los diversos parámetros químico-físicos. Para ello, hay que enfocarse en la descripción matemáticamente desarrollada a partir de la aproximación matricial

de intercambio conformacional y de relajación completa (ecuación 1.2.7). En el trabajo donde se desarrolló esta aproximación, los autores analizaron los factores que influyen en los cambios de las intensidades de los protones del ligando utilizando dos modelos ideales (Jayalakshmi and Krishna 2002). Este estudio mostró la acentuada dependencia entre las intensidades de las señales del ligando y los tiempos de saturación, la posición relativa de los protones saturados de la proteína con respecto a los protones del ligando, la conformación de la interface proteína-ligando, la proporción entre especies, los tiempos de correlación rotacionales de ambas moléculas y la cinética del complejo formado reversiblemente.

En nuestro estudio, los efectos causados por los tiempos de correlación y las cinéticas moleculares van ser analizados en sistemas reales. Primero, se va a analizar la influencia de estos factores sobre el STD-NMR promedio para los tres complejos moleculares presentados en el apartado 5.1.2. Posteriormente, los resultados se utilizarán para explicar el comportamiento de los seis factores R-NOE frente a los parámetros químico-físicos.

Las intensidades medias de STD-NMR fueron obtenidas por ligando teniendo en cuenta diferentes tiempos de correlación rotacionales, velocidades de disociación y constantes de asociación (Figura 5.1.2). Los valores de STD-NMR promedios y las desviaciones estándar se calcularon a partir de los valores de STD-NMR de los protones de cada ligando. Cada parámetro fue explorado entorno a los valores que mejor se ajustaron a los resultados experimentales (Tabla 5.1.1). Además, se consideró que los sistemas presentaban condiciones débiles de asociación ($k_{\text{off}} > 1 \text{ Hz}$) (Meyer and Peters 2003, Lepre, Moore et al. 2004) y una eficiente difusión de spin ($10^{-7} \text{ s} \leq \tau_c \leq 10^{-9} \text{ s}$).

A partir del análisis del comportamiento del STD-NMR promedio en cada sistema se pudo conocer que la saturación promedio muestra una dependencia exponencial en todos los casos. Este comportamiento indica que el STD-NMR está gobernado en los tres sistemas por la matriz dinámica ($\mathbf{D}=\mathbf{R}+\mathbf{K}$)

(ecuaciones 1.2.8 y 1.2.9). El tamaño y la forma globular de las tres proteínas favorecen una difusión de spin eficiente. Con estas características físicas, la transferencia directa de saturación desde los protones inicialmente saturados de las proteínas hacia los protones de los ligandos enlazados (matriz **Q** de la ecuación 1.2.7) no es predominante con un tiempo de saturación de ~ 2 s. Por consiguiente, solo la transferencia indirecta y el intercambio entre estados (matriz **D**) gobiernan las intensidades de STD-NMR.

En los paneles A, B y C de la Figura 5.1.2 se observan los valores de STD-NMR de la kifunesina, salacinol y epotilona-A frente a los tiempos de correlación rotacionales. Al comparar los tres paneles se observa que las desviaciones de los datos calculados en la epotilona-A son menores que en la kifunesina y el salacinol. La diferencia es debida a las características químico-físicas de los tres complejos moleculares. Si comparamos los valores de la Tabla 5.1.1 encontramos que la kifunesina y el salacinol tienen un $\tau_c^p \approx 10^{-7}$ s y una $K_a \approx 10^4$ M $^{-1}$, mientras que en la epotilona-A son $\tau_c^p \approx 10^{-8}$ s y $K_a \approx 10^5$ M $^{-1}$. El aumento de la constante de asociación aumenta el tiempo de residencia de la epotilona-A en el sitio de unión. En el estado enlazado, el tiempo de correlación del ligando se iguala al tiempo de correlación de la proteína. Por ser la β -tubulina de menor tamaño, su frecuencia de rotación disminuye asemejándose más a la frecuencia de rotación del spin nuclear ($\sim 10^{-9}$ s) mejorando la eficiencia de la difusión de espín. En conjunto, los cambios favorecieron los procesos de relajación en el complejo formado por la epotilona-A, afectando el STD-NMR promedio y la diferencia de intensidades entre los protones del ligando.

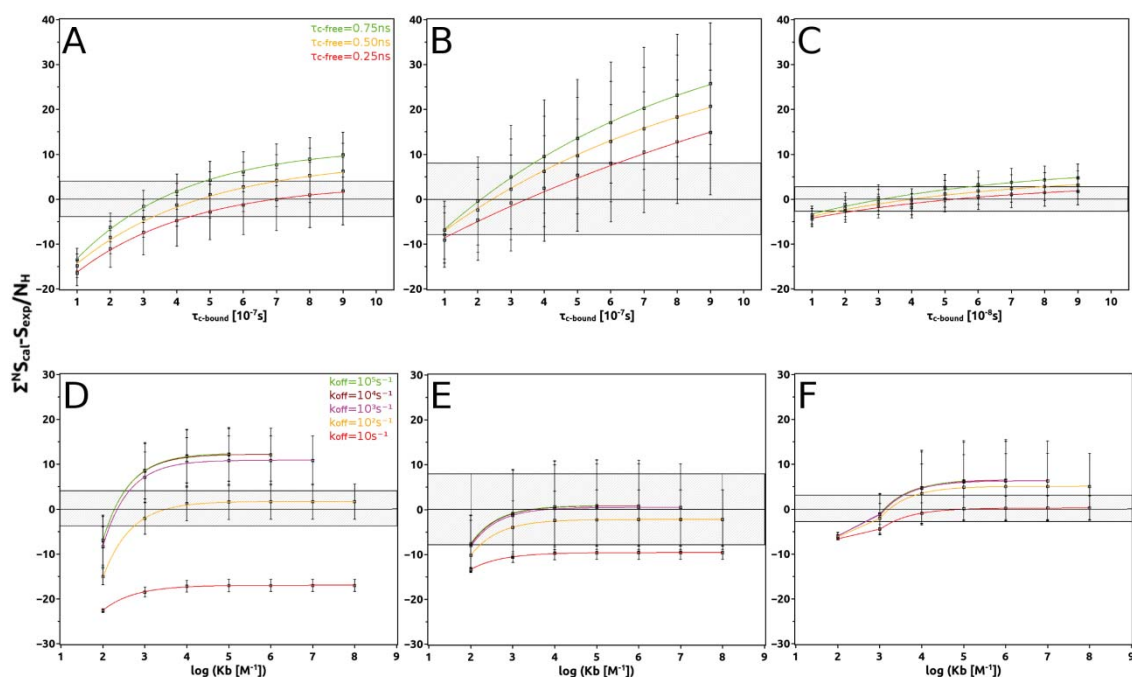


Figura 5.1.2. Representación de la variación del STD-NMR promedio experimental y calculado del complejo kifunesina-dGMII (A y D), salacinol-dGMII (B y E) y epotilona A- α,β -heterodímero de tubulina (C y F). El promedio de cada complejo molecular se calculó a partir de las intensidades de los protones del ligando saturados tomando como referencia el promedio experimental. El área sombreada y las barras de errores representan la desviación cuadrática media de los datos experimentales y teóricos respectivamente. N_H es el número de protones saturados por ligando.

Pero, como se comentó al principio del apartado, las intensidades promedio de cada compuesto aumentan con el incremento de los tiempos de correlación de los ligandos unidos y libres. A primera vista, este resultado parece contradecir el análisis anterior. Aunque, si tenemos en cuenta que los tres sistemas cumplen con una difusión de spin eficiente y que los valores explorados en cada sistema se mantienen con el mismo orden de magnitud, podemos considerar que los cambios del τ_c solo inducen un aumento del tamaño teórico de las moléculas y no modifican susceptiblemente la eficiencia de la relajación.

Si el aumento del tamaño de las moléculas favorece la propagación de la saturación por el incremento de la red dipolar de las moléculas, los tres complejos moleculares van a ser capaces de relajar lo suficientemente rápido su magnetización, como para causar el incremento de la saturación de los protones del ligando sin que la autorelajación del ligando reduzca sus intensidades.

En la Figura 5.1.2, los paneles D, E y F muestran la influencia de la constante de asociación sobre la saturación. Entender tal relación nos ayudará a desenmarañar los efectos causados por la relajación, las poblaciones moleculares y las velocidades de intercambio sobre la transferencia de saturación. Como ya se había anotado, en los tres sistemas, las intensidades promedio crecen exponencialmente con el aumento de la constante de asociación hasta alcanzar una meseta entre $\sim 10^4 \text{ M}^{-1}$ y 10^8 M^{-1} . Un comportamiento similar fue encontrado en la descripción teórica de un complejo molecular simplificado formado por una proteína con 6 protones y un ligando con 5 protones (Jayalakshmi and Krishna 2002).

Es conocido que el rango de $K_a \leq 10^7 \text{ M}^{-1}$ conlleva un intercambio rápido entre el estado libre y unido, ocurriendo este proceso mucho antes que la relajación. Cuando esto ocurre, la matriz de relajación (**R**) se comporta como una perturbación menor de la matriz cinética (**K**). Así, el tratamiento de la teoría perturbacional se aplica a la ecuación 1.2.8 (Krishna and Moseley 2002) y se obtiene:

$$\begin{bmatrix} I_L + I_{L'} \\ I_{E1} + I_{E1'} \end{bmatrix} = \begin{bmatrix} I_{L0} + I_{L'0} \\ I_{E10} + I_{E1'0} \end{bmatrix} + \{1 - \exp(-\langle R \rangle t)\} \langle R \rangle^{-1} \times \begin{bmatrix} R_{L'E2'} I_{E2'0} \\ R_{E1E2} I_{E20} + R_{E1'E2'} I_{E2'0} \end{bmatrix} \quad 5.1.8$$

Bajo estas condiciones, la matriz de velocidad de relajación promedio generalizada ($\langle R \rangle$) es expresada como la suma de los productos de las velocidades de las matrices de relajación y las poblaciones correspondientes al

estado libre y unido (Krishna, Goldstein et al. 1980, Moseley, Curto et al. 1995). Al sustituir los términos de la relajación en ambos estados se termina por obtener:

$$\langle R \rangle = R_l P_l + R_c P_c = \begin{bmatrix} p_L R_L + p_{L'} R_{L'} & p_{E'} R_{L'E1'} \\ p_{L'} R_{E1'L'} & p_E R_{E1} + p_{E'} R_{E1'} \end{bmatrix} \quad 5.1.9$$

Los términos de la diagonal secundaria en la ecuación 5.1.9 describen la relajación intermolecular. Cuando consideramos relaciones ligando/proteína elevadas se cumple que la población del complejo es aproximadamente $p_{E'} \approx 1$. En este caso, los términos $p_{E'} R_{L'E1'}$ y $p_{L'} R_{E1'L'}$ adquieren una relativa importancia, por lo que se puede afirmar que la relajación entre los protones del ligando y la proteína están muy acoplados. Esto nos lleva a concluir que el STD-NMR para $K_a \leq 10^7 \text{ M}^{-1}$ esté gobernado por la velocidad de relajación media generalizada y está particularmente influenciado por la relajación intermolecular.

Aunque el STD-NMR es teóricamente independiente de la constante de asociación para intercambios rápidos, las intensidades disminuyen en los tres sistemas para $K_a \leq 10^4 \text{ M}^{-1}$. Este comportamiento está ligado a la disminución de la población de las moléculas enlazadas. Por debajo de este umbral, la transferencia de saturación desde la proteína hacia el ligando es menor, lo que provoca una reducción de la saturación recibida por los protones del ligando y la disminución de las intensidades de STD-NMR.

Para valores de $K_a > 10^7 \text{ M}^{-1}$, las velocidades de intercambio y de relajación se equiparan y el STD-NMR debe ser descrito obligatoriamente por la ecuación 1.2.8. Sin embargo, nuestros datos no muestran un cambio de comportamiento, así que a $K_a = 10^8 \text{ M}^{-1}$, nuestros tres sistemas reales mantienen el comportamiento descrito por la teoría de la matriz de velocidad de relajación generalizada promedio ($\langle R \rangle$).

A diferencia de K_a , los efectos causados por la k_{off} son evidentes en los cambios sufridos por las intensidades de STD-NMR. El efecto, de carácter cinético, es originado por las diferencias en los tiempos de residencia de los ligandos en el sitio de unión. Observamos en nuestros resultados que cuando disminuye el tiempo de residencia (mayores k_{off}), las intensidades medias y las diferencias entre los protones de los ligandos se incrementan. El tiempo de residencia del ligando unido a la proteína es $\tau_{res}^c = 1/k_{off}$. En este tiempo, los procesos de relajación intermolecular descritos en la ecuación 1.2.8 por los términos $R_{C'[E1'/L]}^T$, $R_{C'[L'/E1]}^{1/2}$ y la autorelajación del ligando descrito por $R_{L'}$ compiten por el desarrollo del STD-NMR (Moseley, Curto et al. 1995, Angulo, Enríquez-Navas et al. 2010). Cuando el tiempo de residencia del ligando enlazado disminuye, el aporte de $R_{L'}$ a la relajación es menor. Al disminuir la autorelajación del ligando, el STD-NMR se incrementa reflejándose tanto en los valores medios como en las diferencias de saturación entre protones de una misma especie.

El análisis exhaustivo del comportamiento del STD-NMR nos permitió entender las causas que influyen en la intensidad de las señales. El análisis del STD-NMR no es simple, por tanto es primordial realizar un análisis cuantitativo profundo para poder interpretar correctamente los datos experimentales. Además, nuestros resultados refuerzan la necesidad de tener extremo cuidado en realizar una interpretación cualitativa considerando que el valor de la intensidad solo está influenciado por la disposición espacial de los protones. Por último, este estudio nos brinda el soporte teórico para poder entender, con mayor claridad, el comportamiento de los factores R-NOE frente a los parámetros cinéticos y los de relajación.

5.1.4 Estudio de la variación de los factores R-NOE con la variación de los parámetros químico-físicos

El primer estudio con los factores R-NOE estará dirigido a analizar si las variaciones matemáticas añadidas son capaces de modificar su robustez y sensibilidad. En este epígrafe deseamos conocer el comportamiento de los factores frente a las modificaciones paramétricas. Por eso, se repetirá el estudio realizado en el epígrafe 5.1.3; pero ahora, teniendo en cuenta los factores R-NOE.

Para estudiar las variaciones de los seis factores (el original y las otras cinco propuestas que hemos hecho en el apartado 5.1.1) frente a los diferentes parámetros químico-físicos, en la Figura 5.1.3 se representaron, los factores R-NOE normalizados (eje y), para tres ligandos diferentes (kifunesina, salacinol y epotilona), en función de los tiempos de correlación rotacionales de los tres complejos moleculares (eje x). Al analizar los resultados se encuentra que todas las funciones están afectadas por los tiempos de correlación rotacionales, independientemente de la modificación matemática añadida. Sin embargo, las funciones R_{tm} y R_{twm} (tercera y cuarta fila contando de arriba abajo) exhiben solo un ligero cambio en los sistemas del salacinol y la epotilona-A (segunda y tercera columna), mientras que en el complejo de kifunesina-dGMII (primera columna) se mantienen invariante cuando el tiempo de correlación de la proteína (τ^p_c) es 10^2 s^{-1} .

Al comparar la tendencia general del conjunto de factores se halla que a los factores R-NOE les afecta más el tamaño de la proteína que el tamaño del ligando. La relación entre el tamaño de la proteína y la intensidad del STD-NMR fueron expuestas en el apartado anterior (5.1.3). En ese apartado, también hablamos de la relación entre el tiempo de correlación de la proteína (τ^p_c), la red dipolar, la relajación interprotónica y la autorelajación, de forma que, pudimos concluir que la relajación interprotónica (ligando-proteína) afecta más a las intensidades del STD-NMR que la autorelajación. La red dipolar de la

proteína es mucho mayor que la del ligando y según los resultados de la Figura 5.1.3, los cambios en las dimensiones de las proteínas favorecen más la propagación de la saturación hacia el ligando. Todo esto se traduce en que los factores R-NOE son más sensibles al tiempo de correlación de la proteína (τ_c^p) que al tiempo de correlación del ligando (τ_c^l). Esta relación explica la cuasi independencia de los factores R_{tm} y R_{tmw} frente al τ_c^l en los complejos de salacinol y epotilona-A.

El análisis de la dependencia paramétrica arrojó que los factores R_e y R_t (primera y segunda fila de la Figura 5.1.3) están más vinculados que el resto de funciones a los tiempos de correlación rotacionales. Ya que, ambos variaron en un 90% sus valores máximos (R_{max}) con el cambio de los tiempos de correlación rotacionales. A su vez, los factores R_{tm} y R_{tmw} (tercera y cuarta fila de la Figura 5.1.3) fueron los más estables, pues ambos presentaron solo un 40% y un 37% de variación máxima (R_{max}) con el cambio de los tiempos de correlación rotacionales. La apreciable diferencia en el comportamiento de estas funciones se debe a la adición de los términos ΔM y ΔM_w (Ecuaciones 5.1.4 y 5.1.5, en el apartado 5.1.1), ya que los dos términos son responsables de reducir las variaciones de la saturación media del ligando. En estos estudios, ΔM y ΔM_w aminoraron los efectos producidos por los cambios en los tamaños teóricos de las moléculas.

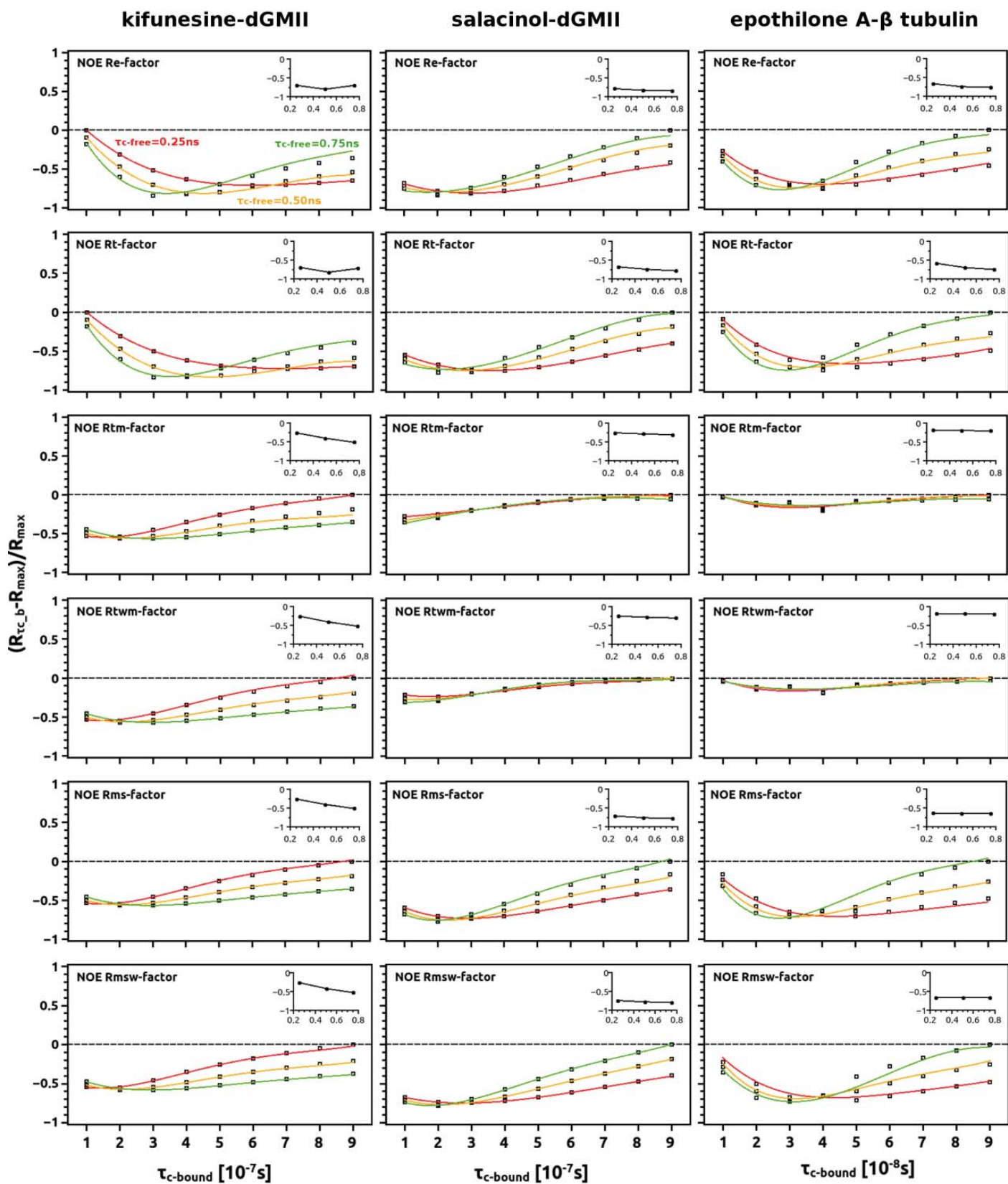


Figura 5.1.3 Promedio normalizado de los factores R-NOE en función de los tiempos de correlación del ligando unido ($\tau_{c-bound}$ o τ_c^p) a distintos valores de tiempo de correlación del ligando libre (τ_{c-free} o τ_c^l). La normalización se realizó tomando como referencia los valores máximos (R_{max}) y ajustándolos a una función exponencial. El tiempo de correlación del ligando libre aparece señalado en la primera gráfica y mantiene el mismo código de colores para todas las demás. En la esquina superior derecha de cada gráfica se encuentran los factores R-NOE en función de los tiempos de correlación de los ligandos libres para los mejores ajustes del STD-NMR experimental promedio.

Otro aspecto a analizar, es el comportamiento de los factores R-NOE en función de la constante de asociación (K_a) y la velocidad de disociación (k_{off}) de los ligandos (Figura 5.1.4). En la figura, podemos observar la respuesta de los seis factores R-NOE (expuestos en el mismo orden que la Figura 5.1.3) para los tres mismos ligandos (kifunesina, salacinol y epotilona) respecto a la constante de asociación (eje X) y a la diferencia normalizada entre los valores de cada factor (R_{K_a}) y su valor máximo (R_{max}) (eje Y). A primera vista, se observa que $\frac{(R_{K_a} - R_{max})}{R_{max}}$, en el complejo de epotilona-A- β -tubulina (tercera columna) resulta ser más sensible que el resto, porque sus variaciones, en este complejo molecular, a distinta velocidad de disociación son mayores. Las pequeñas desviaciones entre las intensidades de STD-NMR calculadas en el apartado anterior (Figura 5.1.2) con la variación de la constante de asociación (K_a) y la velocidad de disociación (k_{off}) han provocado que la desviación máxima de cada factor R-NOE (R_{max}) sea similar a la diferencia ($R_{K_a} - R_{max}$). Por tanto, los pequeños cambios en R_{K_a} inducen cambios relativamente mayores en las diferencias normalizadas en el complejo de epotilona-A- β -tubulina.

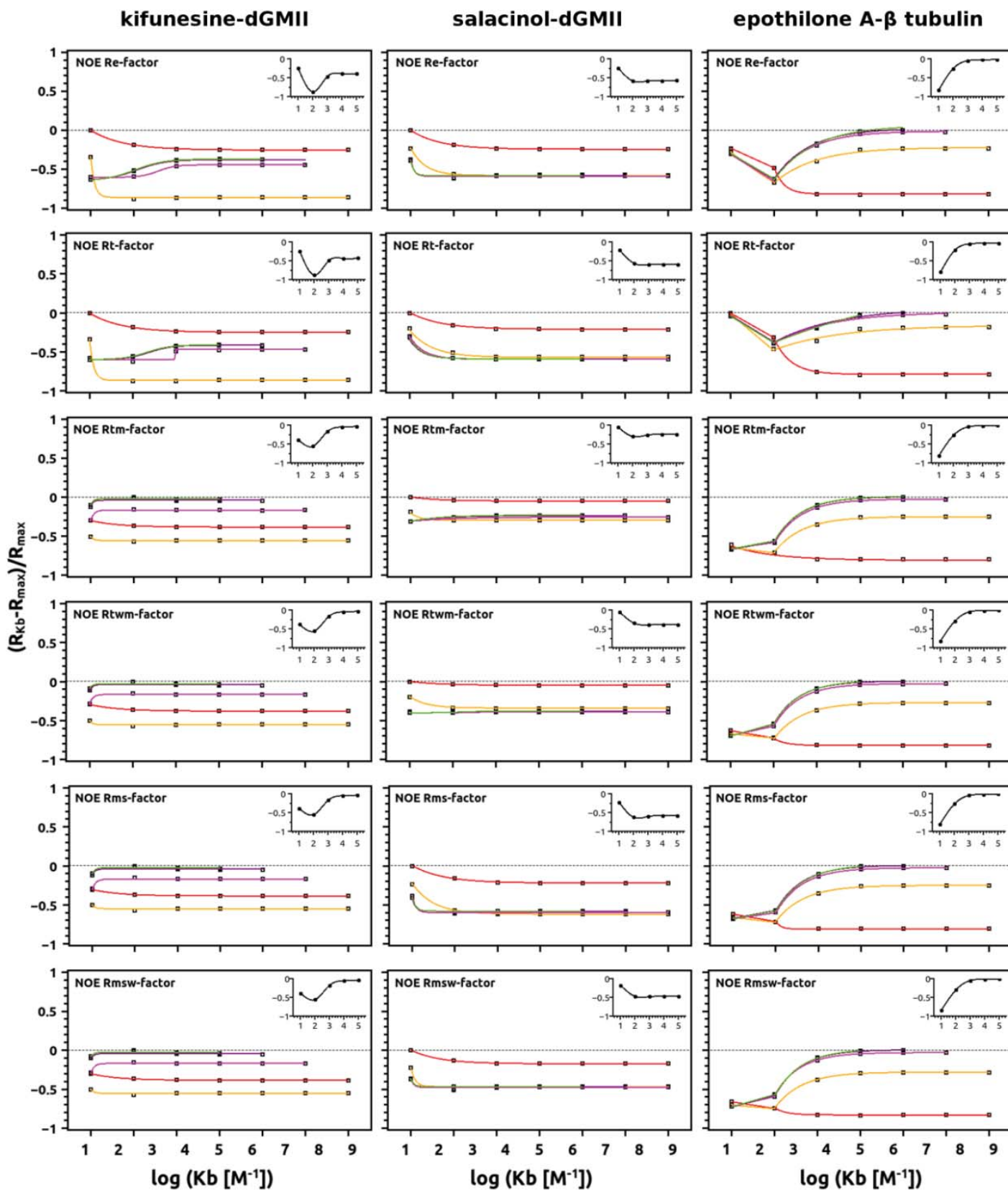


Figura 5.1.4 Factores R-NOE normalizados en función de la constante de equilibrio (K_a) a diferentes valores de velocidad de disociación del ligando (k_{off}) para tres ligandos diferentes (kifunesina, salacinol y epotilona). Los factores fueron normalizados tomando el valor máximo (R_{max}) para cada estudio como referencia. El comportamiento de los factores frente a las velocidades de asociación fue ajustado a una función exponencial y conserva el mismo código de colores para las distintas funciones de la figura. En la esquina superior derecha de cada gráfica se encuentran insertados los factores R-NOE en función de la velocidad de disociación cuando la constante de asociación presenta el mejor ajuste con el STD-NMR experimental promedio.

Anteriormente, se comentó que el STD-NMR está gobernado por la velocidad de relajación media generalizada para $K_a \leq 10^7 \text{ M}^{-1}$ y que para valores menores a 10^4 M^{-1} , la intensidad del STD-NMR decrece porque la transferencia de la saturación desde los protones de la proteína hacia el ligando es apreciablemente menor. Por lo tanto, es esperable que el comportamiento que muestran todos los factores en el rango de $10^4 \leq K_a \leq 10^8 \text{ M}^{-1}$ sea independiente de la constante de asociación. Al analizar solo el comportamiento de las funciones R_{tm} y R_{twm} (segunda y tercera fila), se observa que además no exhiben cambios frente a la pérdida de transferencia de saturación para $K_a \leq 10^4 \text{ M}^{-1}$.

Por el contrario, las modificaciones en la k_{off} , inducen grandes cambios en las funciones, llegando en algunos casos a ser del 90%. Esta tendencia no se cumple en el complejo de salacinol-dGMII (segunda columna). Aquí, tal y como se puede observar, las funciones no parecen estar muy afectadas por la k_{off} (el comportamiento de cada factor R-NOE a diferentes k_{off} es similar) e incluso se produce un solapamiento a determinadas condiciones. Al igual que en el anterior estudio, los factores R_e y R_t (primera y segunda fila) son los más dependientes a las variaciones de K_a y k_{off} . Mientras el resto de los factores (R_{tm} , R_{twm} , R_{ms} y R_{msw}) (tercera, cuarta, quinta y sexta columna) conserva el mismo comportamiento a diferentes K_a y k_{off} , exhibiendo una variación

equivalente al 50% del valor máximo alcanzado en cada estudio por cada factor R-NOE (R_{\max}).

5.1.5 Análisis cuantitativo del efecto de los parámetros químico-físicos en los factores R-NOE

El análisis cualitativo del comportamiento de las seis funciones permite predecir la respuesta de los factores R-NOE en el ajuste de los parámetros (τ_c^p , τ_c^l , K_a y k_{off}) en la ecuación del CORCEMA (ecuación 1.2.7). Aunque esta información es de utilidad, preferimos realizar un análisis cuantitativo para conocer sus efectos cuando se implementen en un procedimiento de resolución estructural. Con este fin, hemos realizado el análisis estadístico de las variaciones de los factores R-NOE para conocer si los términos añadidos (específicamente ΔM y ΔM_w) reducen la dependencia paramétrica significativamente.

En un primer estudio, a partir de los datos de la Figura 5.1.3 y Figura 5.1.4 (apartado anterior) se han obtenido los valores de las variaciones máximas de los factores R-NOE para los tres complejos moleculares (kifunesina-dGMII, salacinol-dGMII y epotilona-A- β -tubulina) y se han representado en un diagrama de caja y bigote (Figura 5.1.5). Hemos utilizado esta representación porque nos permite, de manera visual, conocer la dispersión y simetría de los datos de los tres sistemas extraídos de Figura 5.1.3 y Figura 5.1.4.

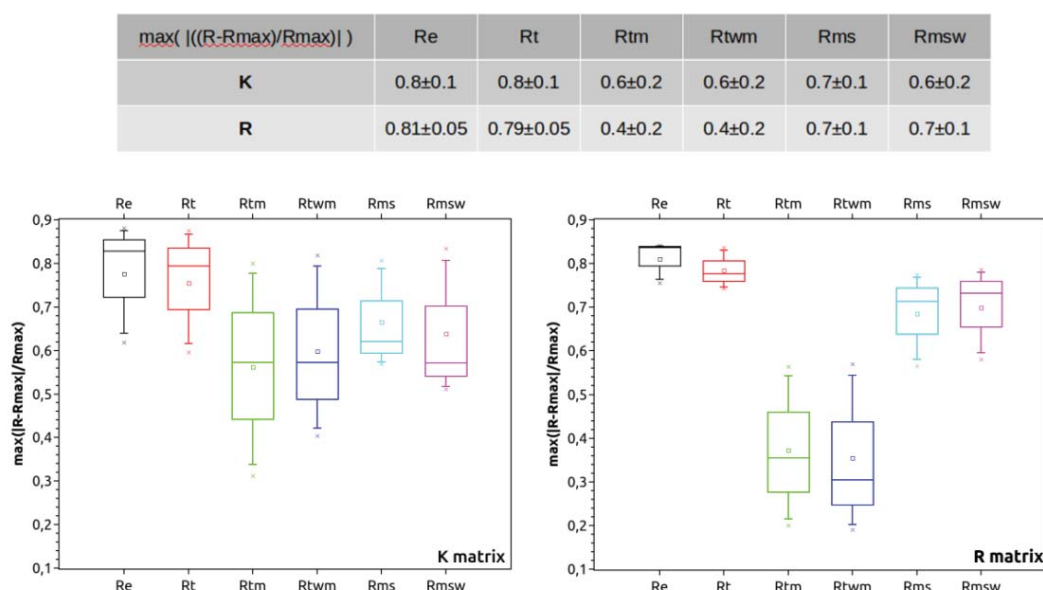


Figura 5.1.5. Diagrama de caja y bigote de la variación máxima normalizada de los factores R-NOE en función de los parámetros de las matrices **K** (ecuación 1.2.9) y **R** (ecuación 1.2.8). En la parte superior se encuentra una tabla con la media de los valores de los seis factores R-NOE y sus desviaciones estándar.

En la imagen se observa que existe una menor dependencia de los factores R_{tm} y R_{twm} , mientras que los factores R_e y R_t modifican en gran medida sus valores a causa de los parámetros de las matrices de relajación (ecuación 1.2.8) y cinética (ecuación 1.2.9). Al analizar a R_{ms} y R_{msw} , se puede ver que las modificaciones matemáticas realizadas en ambos factores para $i=50$ también logran reducir los efectos de la matriz de relajación y la cinética. Aunque la tendencia de R_{tm} , R_{twm} , R_{ms} y R_{msw} es a disminuir las variaciones frente a los parámetros cinéticos y de relajación, con los datos de la figura no podemos asegurar que la reducción en la sensibilidad paramétrica es significativa.

Para estimar si las diferencias mostradas por los factores R-NOE son estadísticamente significativas se utilizará una prueba estadística. En estudios similares, por lo general se admite la presencia de un error de tipo I de un 5%,

que es lo mismo que considerar un intervalo de confianza del 95% o un nivel de significancia de 0,05. El error tipo I es el error que se comete al detectar diferencias significativas entre dos muestras cuando no existen (“falso positivo”) y la probabilidad de tomar esta decisión viene definida por el nivel de significancia (α , en este caso en particular $\alpha=0.05$). Aunque un 5% de error es asumible en una prueba estadística, este puede incrementarse si el estudio se realiza más de una vez al comparar más de dos medias muestrales. En nuestro estudio, vamos a comparar las respuestas de los seis factores R-NOE, así que es importante realizar el análisis con un estadístico que no aumente el error tipo I con el aumento de las muestras. Por lo que en nuestra evaluación se ha utilizado un Análisis de Varianzas (ANOVA).

La ANOVA es una técnica estadística que permite determinar si los factores R-NOE muestran diferencias significativas o por el contrario, se puede suponer que no existe ninguna diferencia entre ellos. La ventaja de la prueba de ANOVA se encuentra en el control del error tipo I, ya que no se incrementa al aumentar el total de grupos a comparar. La ANOVA no está libre de limitaciones; para poder utilizarla, las muestras tiene que ser independientes, mostrar una distribución poblacional normal y cumplir con una homocedasticidad entre muestras (la varianza entre las poblaciones es igual).

Al plantearnos el ensayo estadístico, dividimos los datos en seis grupos diferentes, uno por cada factor R-NOE. Cada grupo contiene un tamaño población igual a tres porque los datos provienen del estudio de tres complejos moleculares. En poblaciones tan pequeñas, la posibilidad de que las tres condiciones se cumplan no es asumible. Es por ello que en el anexo 9.2 se comprueba el cumplimiento de las tres condiciones. Para saber si nuestros datos son aptos para un análisis ANOVA, tres pruebas estadísticas previas fueron realizadas (anexo 9.2). Dos de ellas estaban dirigidas a comprobar la homocedasticidad y la última para evaluar la normalidad de la población. Las tres pruebas confirmaron, con un error tipo I del 5%, que todos nuestros datos

cumplen con las condiciones necesarias para poder proseguir con el estudio de la ANOVA.

Posteriormente a la confirmación, definimos nuestras hipótesis. Así que se consideró como hipótesis nula que los efectos que causan los parámetros afectan por el igual a todos los factores R-NOE. Además una hipótesis alternativa es planteada y con ella se define que los efectos de los parámetros cinéticos y de relajación se reducen en algunos de los factores R-NOE. Para realizar la prueba se van a considerar dos grupos, en el primero se estudiará el comportamiento de los factores R-NOE frente a los cambios en K_a y k_{off} (cambios en la matriz cinética (\mathbf{K})), mientras que en segundo grupo se estudian las variaciones en los factores debidas a τ_c^p y τ_c^l (cambios en la matriz de relajación (\mathbf{R})). Por eso, dos pruebas ANOVA de una cola fueron realizadas. En nuestro estudio consideramos la prueba de una cola porque esperamos que los efectos se reduzcan o queden igual entre los seis factores R-NOE.

La técnica fundamental en la ANOVA consiste en la separación de la suma de cuadrados total (SS_{total} , “*sum of squares*”) en dos componentes ($SS_{intergrupo}$ y $SS_{intragrupo}$). La suma de cuadrados representa una medida de variación o desviación con respecto a la media. En el análisis de varianza (ANOVA), la suma total de los cuadrados ayuda a expresar la variación total como la variación que se atribuye a los diferentes factores R-NOE (análisis intergrupo, $SS_{intergrupo}$) y a la variación debida a los tres sistemas moleculares estudiados (kifunesina-dGMII, salacinol-dGMII y epotilona-A- β -tubulina)(análisis intragrupo, $SS_{intragrupo}$). La comparación entre las dos variaciones ($SS_{intergrupo}$ y $SS_{intragrupo}$) permite identificar los cambios en el comportamiento de los factores R-NOE. En la prueba de significancia, se utiliza la distribución F para cuantificar la relación entre la variación debida a los factores ($SS_{intergrupo}$) y la debida a los sistemas moleculares ($SS_{intragrupo}$). Una vez se han calculado las sumas de cuadrados, se considera el nivel de significancia deseado (α) y se calcula el valor p. El valor p (p-valor) es la probabilidad de obtener un resultado al menos tan extremo como el que realmente se ha obtenido a partir del valor del estadístico utilizado. Al

obtener el valor p y haber definido el nivel de significancia deseado (α) se puede conocer si los resultados presentan diferencias significativas o no (permite escoger la hipótesis alternativa o la hipótesis nula como ciertas). Finalmente se procede a representar los datos de las medias cuadráticas, los grados de libertad y la F (Tabla 5.1.2).

Tabla 5.1.2. Resultado del estudio estadístico ANOVA de una cola.

Diferencias máximas por cambios cinéticos					
Fuente	Suma de cuadrados ($SS_{total}=SS_{inter}+SS_{intra}$)	Grados de libertad (df)	Media cuadrática ($SS_{inter \text{ o } intra}/df$)	Valor F $\left(\frac{SS_{inter}/df_{inter}}{SS_{intra}/df_{intra}}\right)$	Valor p
Intergrupos	0,1087	5	0,0217	0,6940	0,6379
Intragrupos	0,3760	12	0,0313		
Total	0,4848	17			
Al nivel de significancia de 0.05, las medias poblacionales no son significativamente diferentes					
Diferencias máximas por cambios en la relajación					
Intergrupos	0,6161	5	0,1232	7,4764	0,0021
Intragrupos	0,1978	12	0,0165		
Total	0,8139	17			
Al nivel de significancia de 0.05, las medias poblacionales son significativamente diferentes					

En nuestro estudio, establecimos como intervalo de confianza $\alpha=0,05$. Por lo tanto, si el valor p calculado a partir de la distribución F es menor que 0,05 el comportamiento de los factores R-NOE será significativamente diferentes. Teniendo esto en cuenta, calculamos el valor p en los dos estudios de ANOVA para poder analizar los resultados estadísticos. A partir de los resultados concluimos que con una $F(5,12)=0.694$ y $p=0.64$, el comportamiento de los factores R-NOE frente a la influencia de la matriz **K** es estadísticamente el

mismo. Sin importar que tipo modificación matemática se añadiera al factor R-NOE (ΔM y ΔM_w), el efecto de la k_{off} sobre los cuatro factores R-NOE no se llegó a reducir significativamente.

Por otra parte, la prueba ANOVA dio que las variaciones de los factores R-NOE con el cambio de los tiempos de correlación sí son estadísticamente significativas ($F(5,12)=7.476$, $p=0.002$). Por lo tanto, las adiciones matemáticas (ΔM y ΔM_w) sí rectifican la desviación de los factores R-NOE al cambiar los tiempos de correlación rotacionales.

Hasta ahora, el estudio estadístico ha permitido conocer si entre los factores R-NOE hay un comportamiento diferenciado frente a los cambios de τ^p_c , τ^l_c , K_a y k_{off} . Los resultados también indican que los valores de algunos de los factores R-NOE varían menos a causa de modificación en los parámetros químico-físicos, lo que implica una menor dependencia hacia estos parámetros y una muestra de robustez por parte de los factores. Pero, con la ANOVA no se pueden identificar cuáles son los factores que se diferencian estadísticamente del resto.

Así que una vez se ha determinado la existencia de una diferencia entre los factores, podemos aplicar la prueba de rango post-hoc y las comparaciones múltiples por parejas. Este tipo de pruebas permite determinar cuáles son los factores R-NOE con un comportamiento estadísticamente diferenciado y que, por tanto, muestren una dependencia paramétrica reducida. En este paso, la prueba de Tukey (Lomax and Hahs-Vaughn 2013) es usualmente utilizando para identificar subconjuntos homogéneos de medias que no difieran entre sí. La prueba estadística de Tukey está basada en las mismas suposiciones que la ANOVA y por este motivo, se suelen aplicar conjuntamente.

La prueba de Tukey (HSD, "*Honestly-significant different*") es una prueba estadística de comparaciones múltiples que se utilizará para comparar las variaciones de los factores R-NOE frente a los cambios de las constantes químico-físicas. En el estudio se opta por aplicar esta prueba porque con ella

podremos especificar el o los factores que redujeron el efecto de los parámetros químico-físicos. Al igual que en el estudio de la ANOVA, en el que se utilizaba el estadístico F, aquí se utilizará el estadístico Q. Q como en la mayoría de los estadísticos es una función que depende de los datos en estudios, está diseñada para ser sensible a la hipótesis alternativa (típicamente, su magnitud tiende a ser mayor cuando la hipótesis alternativa es verdadera) y en su caso particular, cumple con una distribución normal (distribución gaussiana).

Así que se realizó una comparación estadística por cada pareja de factores R-NOE, para evaluar las diferencias que muestran entre ellos, frente a cambios paramétricos en la matriz cinética y en la matriz de relajación. En el estudio de Tukey, la hipótesis nula y la hipótesis alternativa se plantearon con las mismas distinciones que se tuvieron con el estudio de la ANOVA. Las comparaciones por parejas y los resultados del test de Tukey son mostrados en la Tabla 5.1.3. En ella se muestran los valores calculados del estadístico Q, la probabilidad de tener el resultado (valor p) y la significación estadística de la comparación a partir de los datos de Q y el valor p. Las comparaciones por parejas de factores R-NOE que muestran un valor p menor que el intervalo de confianza considerado ($\alpha=0.05$) fueron subrayados y resaltados en color rojo. La prueba corrobora que los seis factores R-NOE mantienen el mismo comportamiento frente a las modificaciones en los parámetros de la matriz **K**. Sin embargo, en la comparación frente a las modificaciones en los parámetros de la matriz **R**, la hipótesis alternativa fue dada como cierta cuatro veces al ser el p valor menor que el nivel de significancia definido ($\text{valor-p} < \alpha$). Como se puede observar en la tabla, a través de esta prueba se constata que las respuestas brindadas por los factores R_{tm} y R_{twm} son estadísticamente diferentes a las dadas por R_e y R_t . Mientras que, el resto de las comparaciones no mostraron diferencias significativas según la prueba de Tukey.

Tabla 5.1.3. Prueba estadística post hoc de Tukey

	Matriz Cinética			Matriz de Relajación		
	Estadístico Q	p- valor	Conclusión	Estadístico Q	p- valor	Conclusión
R_e vs. R_t	0.2025	0.8999	Insignificante	0.3454	0.8999	Insignificante
R_e vs. R_{tm}	2.0965	0.6638	Insignificante	5.9113	0.0125	* $p < 0.05$
R_e vs. R_{twm}	1.7374	0.7949	Insignificante	6.1467	0.0095	** $p < 0.01$
R_e vs. R_{ms}	1.0815	0.8999	Insignificante	1.7055	0.8066	Insignificante
R_e vs. R_{msw}	1.3413	0.8999	Insignificante	1.5067	0.8791	Insignificante
R_t vs. R_{tm}	1.8940	0.7377	Insignificante	5.5658	0.0189	* $p < 0.05$
R_t vs. R_{twm}	1.5350	0.8688	Insignificante	5.8013	0.0143	* $p < 0.05$
R_t vs. R_{ms}	0.8790	0.8999	Insignificante	1.3601	0.8999	Insignificante
R_t vs. R_{msw}	1.1388	0.8999	Insignificante	1.1613	0.8999	Insignificante
R_{tm} vs. R_{twm}	0.3591	0.8999	Insignificante	0.2355	0.8999	Insignificante
R_{tm} vs. R_{ms}	1.0150	0.8999	Insignificante	4.2058	0.0945	Insignificante
R_{tm} vs. R_{msw}	0.7552	0.8999	Insignificante	4.4045	0.0752	Insignificante
R_{twm} vs. R_{ms}	0.6559	0.8999	Insignificante	4.4412	0.0721	Insignificante
R_{twm} vs. R_{msw}	0.3961	0.8999	Insignificante	4.6400	0.0570	Insignificante
R_{ms} vs. R_{msw}	0.2598	0.8999	Insignificante	0.1988	0.8999	Insignificante

*Nivel de significancia de 0.05

**Nivel de significancia de 0.01

Al recopilar todos los resultados obtenidos en este epígrafe, pudimos dividir las funciones en tres subgrupos por su variabilidad. El primer subgrupo está formado por los factores R_{tm} y R_{twm} , pues son las funciones menos dependientes con una variabilidad máxima de solo un 40 ± 20 %. El segundo subgrupo está formado por R_e y R_t . Éstos son los factores más dependientes a los parámetros cinéticos y de relajación. El último subgrupo está formado por los factores R_{ms} y R_{msw} debido a que para $i=50$ sus respuestas son mediocres al no presentar diferencias significativas con respecto al resto de los factores.

Finalmente, a partir de los estudios estadísticos realizados en el actual apartado se llegó a demostrar que los factores R_{tm} y R_{twm} mejoraron su robustez al presentar una reducción estadísticamente significativa de la variabilidad de sus valores provocada por los cambios en los paramétrico τ_c^p y τ_c^l . Los factores R_{tm} y R_{twm} solo se diferencian de R_l en los términos ΔM o ΔM_w , así que la mejora significativa de la robustez se debe estrictamente a la adición de estos términos.

5.1.6 Análisis de la variación de los factores R-NOE con el refinamiento estructural

Los factores R-NOE han sido diseñados para evaluar la estructura de un complejo molecular mediante un proceso predictivo basado en datos de STD-NMR. Antes de poder implementarlos, nos estamos cerciorando del correcto funcionamiento de todos ellos mediante diversos estudios. En el primer análisis (apartados 5.1.4 y 5.1.5), caracterizamos la relación entre los factores R-NOE y los parámetros químico-físicos, así que en los siguientes dos apartados, se estudiarán los cambios en los factores inducidos por las coordenadas de los complejos moleculares. Estos dos estudios nos definirán la utilidad de los factores R-NOE y nos brindarán información acerca de la sensibilidad de estos en la identificación de las estructuras descritas por los datos de STD-NMR.

Las metodologías de elucidación de estructuras mediante datos de STD-NMR conllevan dos pasos previos: uno de exploración conformacional y otro de refinamiento estructural. Las técnicas de modelado molecular utilizadas en ambos procesos son diversas, entre ellas podemos nombrar: la mecánica molecular para la optimización estructural (Canales 2011, Field, Pera et al. 2012, Antúñez-Mojica, Rodríguez-Salarichs et al. 2016), la dinámica molecular para el relajamiento de las interacciones (Matesanz, Rodríguez-Salarichs et al. 2011, Szczepina 2011) y la exploración conformacional mediante algoritmos

genéticos o de simulación por Monte Carlos (Canales, Nieto et al. 2014, Jantsch, Nieto et al. 2016). Estas técnicas, aún con las diferencias que tienen en su fundamento físico, tienen un punto en común, pues todas necesitan un campo de fuerzas para evaluar las distancias entre los átomos. Normalmente, se recurre a campos de fuerzas convencionales como OPLS (Jorgensen and Tiradorives 1988, Jorgensen, Maxwell et al. 1996), AMBER (Jorgensen, Maxwell et al. 1996), CHARMM (Vanommeslaeghe, Hatcher et al. 2010), etcétera; por tanto, los datos de STD-NMR quedan relegados a la evaluación de la estructura final, posterior a los procesos de construcción y refinamiento molecular. Por consiguiente, el uso de los datos de STD-NMR en la evaluación estructural es pasivo al no utilizarse en los procesos de exploración espacial.

Por primera vez en el 2004, una metodología híbrida (SICO) fue capaz de refinar la estructura de un complejo molecular mediante la implementación de una función de energía basada en el factor R_e (Jayalakshmi and Rama Krishna 2004). El procedimiento SICO demostró que el uso activo del factor R_e en la evaluación de estructuras de complejos moleculares es posible. Para conocer si los factores R-NOE son sensibles al cambio conformacional y si es posible su uso en metodologías de refinamiento, vamos a analizar el comportamiento de los factores R-NOE como funciones de energía en la exploración conformacional de las estructuras de la kifuensina, salacinol y epotilona-A enlazadas a sus dianas.

Las coordenadas 3D de las tres estructuras cristalográficas fueron tomadas como estructuras de partida. El espacio conformacional de cada ligando fue explorado utilizando la función de energía OPLS2005, la diferencia entre las intensidades promedio del STD-NMR calculado y experimental, y los seis factores R-NOE. Los ángulos torsionales se rotaron 360° en pasos de 5° y en cada paso, se realizó un posterior refinamiento estructural para eliminar los choques estéricos y las estructuras poco probables.

Los valores resultantes de las tres exploraciones conformacionales fueron normalizados y representados en la Figura 5.1.6. Debido a la simplicidad

estructural de la molécula de kifunesina (Figura 5.1.6-A), solo el ángulo torsional del metanol unido al anillo fue rotado. Para los otros dos ligandos (salacinol, Figura 5.1.6-B y epotilona-A, Figura 5.1.6-C), dos ángulos diedros de las zonas más flexibles de las moléculas fueron rotados.

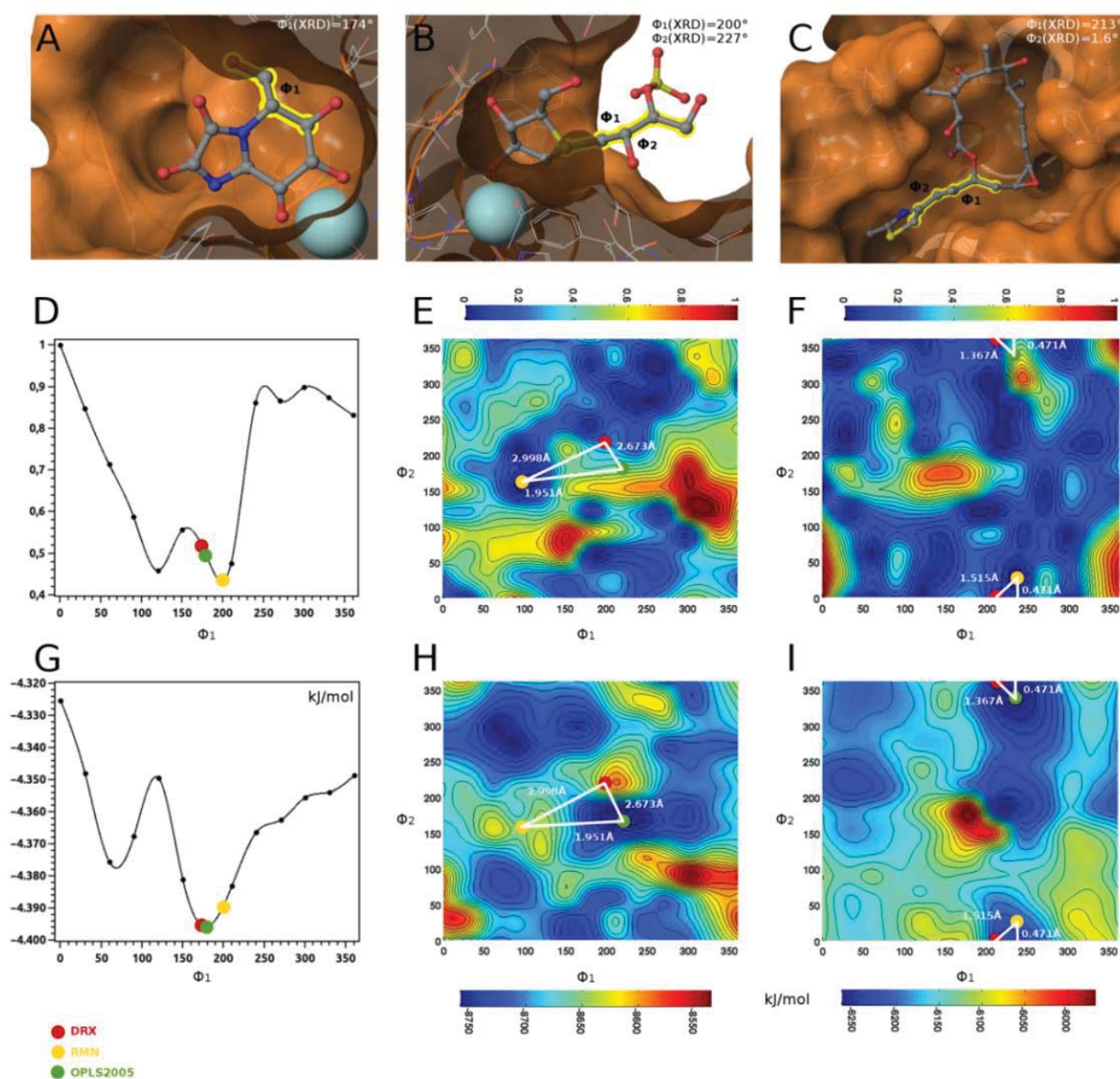


Figura 5.1.6. Exploración conformacional de la kifunesina, el salacinol y la epotilona-A enlazados a sus proteínas dianas. (A-C) Representación de la estructuras cristalográficas de la kifunesina unida a dGMII (A), salacinol unido a dGMII (B) y la epotilona-A unida al complejo T₂R (C). Los átomos que forman los planos diedros están resaltados en amarillo. (D-F) Representación del STD-NMR promedio normalizado en la exploración conformacional de la kifunesina (D), el salacinol (E) y la epotilona-A (F). (G-I) Representación de la energía torsional descrita por el campo de fuerzas OPLS en la exploración conformacional de la kifunesina (G), el salacinol (H) y la epotilona-A (I). Los puntos señalan las conformaciones encontradas por los métodos de DRX (rojo), las diferencias entre las intensidades experimentales y calculadas de STD-NMR (amarillo) y el campo de fuerzas OPLS2005 (verde). Además, se muestran los valores de RMSD entre las conformaciones.

En la exploración rotacional del ángulo diedro del complejo de kifunesina-dGMII (Figura 5.1.6-A), el campo de fuerzas teórico (Figura 5.1.6-G) y los valores de STD-NMR (Figura 5.1.6-D) muestran un comportamiento similar. El estudio conformacional basado en el STD-NMR describe dos mínimos de similar intensidad a 120° y 200° (Figura 5.1.6, panel D). A su vez, la superficie del campo de fuerzas OPLS2005 (Figura 5.1.6, panel G) tiene un mínimo local intenso alrededor de 170°, y otro menor a 50°. Mientras que, el ángulo torsional de la estructura cristalográfica es 174° (Figura 5.1.6-A). En la comparación del ángulo diedro de la estructura cristalográfica y la estructura del mínimo global del campo de fuerzas, se determinó que el conformero descrito por la estructura cristalográfica es el mismo que el descrito por el mínimo global del campo de fuerzas. En la descripción torsional obtenida a partir de los datos de RMN, encontramos que uno de los mínimos se encuentra desplazado 30° del ángulo torsional de la estructura cristalográfica. Como la diferencia nos induce a pensar que las tres técnicas no describen el mismo conformero; los RMSD de las tres estructuras fueron calculados y sus valores arrojaron diferencias estructurales menores a 0,2 Å. Así que con estas diferencias en la estructura se puede afirmar que las tres técnicas están describiendo el mismo conformero para el complejo de kifunesina-dGMII (Figura 5.1.6-A, D y G).

En los casos más complejos (la exploración de dos ángulos diedros), la descripción conformacional del campo de fuerzas (Figura 5.1.6-H,I) y los datos de RMN (Figura 5.1.6-E,F) hacen visible sus diferencias. En los complejos moleculares de salacinol (Figura 5.1.6-B) y epotilona-A (Figura 5.1.6-C), la mayor parte de las superficies descritas por los datos de RMN difieren de las descritas por el campo de fuerzas (salacinol, Figura 5.1.6-E,H y epotilona-A, Figura 5.1.6-F,I). Las diferencias, en gran parte de las superficies conformacionales, son debidas a las consideraciones que cada modelo realiza.

El campo de fuerzas OPLS2005 es atomístico y describe las interacciones de enlace, angulares, torsionales, electrostáticas y de Van der Waals entre todos los átomos. Por consiguiente, el campo de fuerzas describe todas las interacciones atómicas de los sistemas. Por otra parte, los datos de STD-NMR resultan de la combinación de los procesos de relajación de la saturación entre los protones de la proteína y el ligando. Como se ha comentado anteriormente, en los ensayos de STD-NMR, el paso de adquisición se realiza entre el primer y quinto segundo. En este rango de tiempo, la información extraída es aportada básicamente por la relajación intermolecular ($R_{C'[E1'/L]}^T$, $R_{C'[L'/E1]}^{1/2}$) y la autorelajación de la molécula de menor tamaño ($R_{L'}$, R_L) (ecuación 1.2.8).

Debido a las diferencias entre ambos métodos (Campo de fuerzas OPLS y acoplamiento entre espines en la técnica de STD-NMR), los comportamientos de los dos métodos se espera que muestren similitudes cuando las interacciones dipolares interprotónicas e internas del ligando sean predominantes. Esta afirmación no es meramente teórica pues, en las superficies de la Figura 5.1.6 se pueden encontrar tres zonas que cumplen con esta condición. En las superficies del complejo de salacinol-dGMII (Figura 5.1.6-E,H), se visualizan regiones superficiales similares entorno a $100^\circ < \Phi_1 < 300^\circ$ y $0^\circ < \Phi_2 < 100^\circ$; mientras que en el complejo de epotilona-A- β -tubulina (Figura 5.1.6-F,I) se observan dos zonas, la primera comprendida entre $200^\circ < \Phi_1 < 300^\circ$ y $0^\circ < \Phi_2 < 100^\circ$ y la segunda entre $0^\circ < \Phi_1 < 100^\circ$ y $200^\circ < \Phi_2 < 360^\circ$.

Cuando analizamos las estructuras descritas por los tres métodos, para los complejos de salacinol (Figura 5.1.6-B, E y H) y epotilona-A (Figura 5.1.6-C,F,I), encontramos que los ángulos torsionales de las estructuras cristalográficas no coinciden con los conformeros escogidos a partir de los mínimos de energía o los de RMN. Incluso, la estructura cristalográfica del salacinol (Figura 5.1.6-B) se encuentra en un máximo de energía según el campo de fuerzas OPLS (Figura 5.1.6-H). Aunque este resultado parezca inverosímil, la explicación radica en el método de resolución cristalográfico. Primero, los métodos cristalográficos no tienen resolución suficiente para describir los átomos de hidrógenos; además, la estructura obtenida a partir del cristal pueden estar fuera del mínimo de energía por dos razones: la primera, la estructura se resuelve a partir de un cristal y esto puede influenciar sobre la estructura del complejo molecular y la segunda, en el procedimiento de optimización estructural, no se llega a corregir las coordenadas cristalográficas mediante una minimización energética. Estos tres razones son primordialmente los responsables de alejar la estructura cristalográfica del mínimo de energía potencial.

Por el contrario, la información obtenida por técnicas de RMN está influenciada por la estructura de mínima energía. Esto se debe a que el STD-NMR experimental se obtiene a partir de las aportaciones de STD-NMR realizadas por cada estado molecular accesible (conformaciones y orientaciones de baja energía que adopta el complejo molecular). Como conocemos de la física estadística que, los estados de menor energía se encuentran más poblados; las conformaciones y orientaciones de menor energía aportan más información estructural que el resto (Landau and Lifshitz 2013). De esta manera, podemos concluir que los datos de STD-NMR experimentales describen fundamentalmente los acoplamientos entre los protones de la estructura de mínima energía y, por consiguiente, debemos esperar que los mínimos localizados por ambos métodos sean similares siempre que las interacciones dipolares comentadas anteriormente sean predominantes.

Con el fin de conocer si los tres métodos (DRX, STD-NMR y OPLS) describen la misma estructura, se calcularon los RMSD entre los conformeros de mínima energía (Tabla 5.1.4). A partir de los valores obtenidos, podemos concluir que las tres técnicas describen el mismo conformero en los complejos con epotilona-A y kifunesina. Sin embargo, el mismo resultado no se repite en el complejo con el salacinol pues, la estructura descrita por DRX difiere de la estructura descrita por STD-NMR y OPLS. En este caso, los valores elevados de RMSD son suscitados por el cambio en el ángulo torsional Φ_2 . Dicho cambio induce una restructuración en la cadena lateral que modifica sustancialmente la conformación proveniente de la RMN y el campo de fuerzas. Este resultado no se esperaba, porque basados en la evidencia experimental que teníamos, habíamos concluido en el apartado 5.1.2, que tanto los datos de RMN, como las coordenadas cristalográficas, describen la misma estructura para los tres complejos moleculares.

Tabla 5.1.4 Desviación cuadrática media de los conformeros descritos por la DRX, RMN y el modelado molecular.

	DRX		RMN			OPLS2005			
	Φ_1	Φ_2	Φ_1	Φ_2	RMSD ¹	Φ_1	Φ_2	RMSD ¹	RMSD ²
KIF	174	-	200	-	0,154	180	-	0,035	0,166
SAL	200	227	90	160	2,998	225	175	2,673	1,951
EPOA	213	2	240	30	1,515	240	340	1,367	0,471
Estructura de referencia: ¹ coordenadas cristalográficas y ² coordenadas basadas en el STD-NMR									

Sin embargo, al observar el Panel E de la Figura 5.1.6, encontramos que la diferencia entre el STD-NMR calculado y experimental toma un valor mínimo para ángulos torsionales muy cercanos a los descritos por la estructura cristalográfica. El cálculo del RMSD entre las dos conformaciones nos confirma

que las estructuras no se diferencian en más de 2 Å. Así que, este mínimo local ha sido el causante de que los datos teóricos de STD-NMR, obtenidos a partir de las coordenadas cristalográficas en el apartado 5.1.2 (Panel E de la Figura 5.1.1), se asemejaran al experimental. Según estos resultados, concluimos que, los datos de DRX y los de STD-NMR experimentales, difieren en la descripción de la cadena lateral del salacinol en el sitio de unión de dGMII; además, los resultados alertan de la necesidad de utilizar un procedimiento de exploración espacial intensiva haciendo uso activo de los datos de STD-NMR. Además, demuestra fehacientemente que en ausencia de una correcta exploración, se corre el riesgo de escoger estructuras erróneas por falta de información.

Ahora procedemos a evaluar el desarrollo de los factores R-NOE en el análisis conformacional de los tres ligandos. Con este fin, el estudio conformacional se ha realizado utilizando como funciones de energía los 6 factores R-NOE (Figura 5.1.7). Los resultados revelaron que las modificaciones matemáticas presentes en cada factor R-NOE no afectan sus respuestas con el cambio de coordenadas. En la Figura 5.1.7 se puede constatar que solo alteran la profundidad relativa de cada factor y que el comportamiento de todas ellas es el mismo que el del factor R_e .

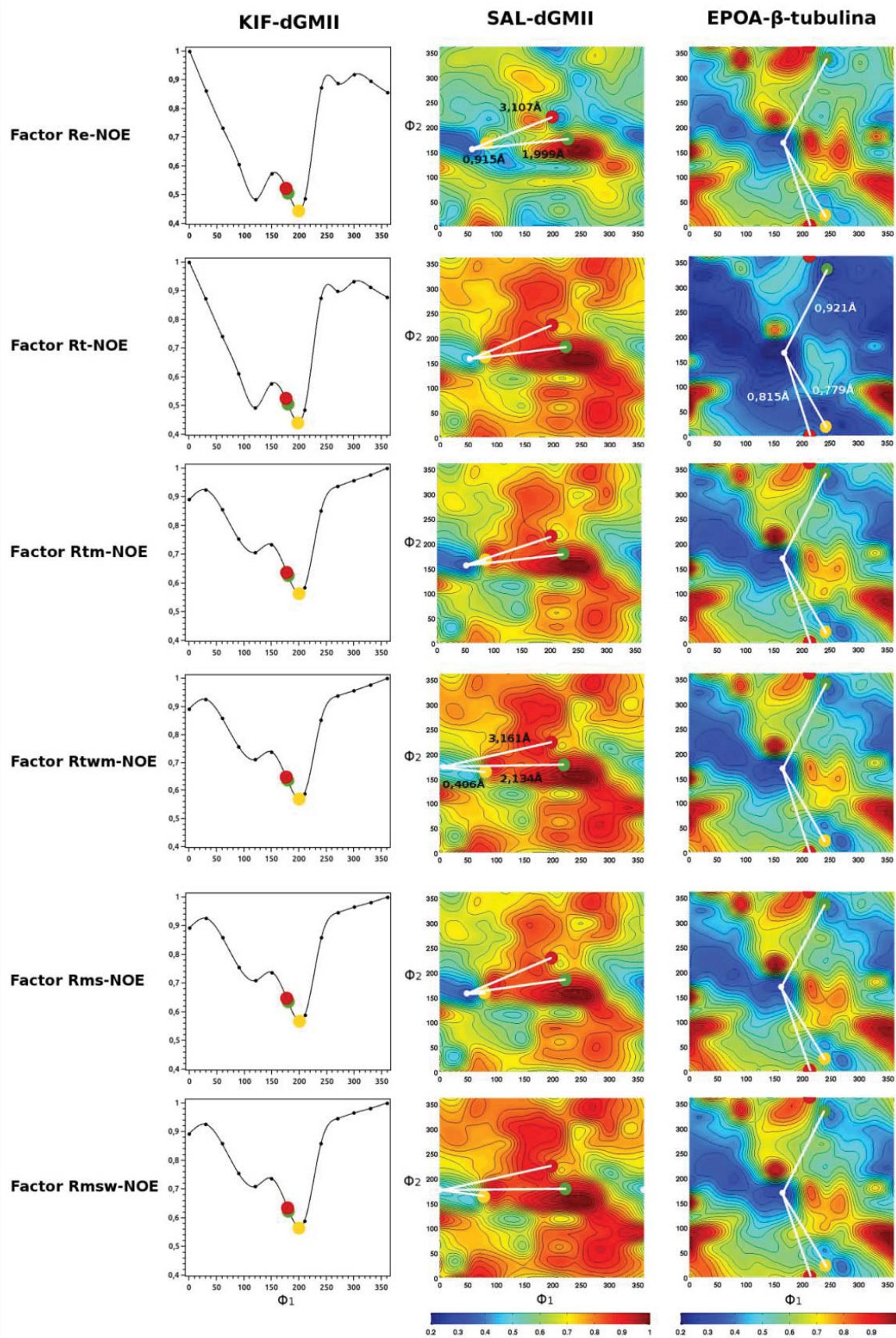


Figura 5.1.7. Representación de las superficies conformacionales de los complejos de kifunesina (columna 1), salacinol (columna 2) y epotilona-A (columna 3) basadas en los factores R-NOE. Los puntos representan las conformaciones encontradas por los métodos de DRX (rojo), diferencias entre las intensidades experimentales y calculadas de STD-NMR (amarillo) y el campo de fuerzas OPLS2005 (verde). Los valores de las funciones han sido normalizados.

Hasta ahora, hemos comparado las superficies de los factores R-NOE, los datos crudos de STD-NMR y las energías. Y curiosamente, se podría decir que el comportamiento de los factores R-NOE puede ser producto de la combinación de los datos de STD-NMR crudos y la función de energía. Al fijarnos en los mínimos globales de las superficies, notamos como los factores minimizan sus valores cuando están cercanos a la localización de estos mínimos. Durante este comportamiento, los factores evidencian su tendencia a buscar estructuras de baja energía que cumplan con la información proporcionada por los datos de STD-NMR. La mimetización del comportamiento de la función de energía por parte de los factores se evidencia en la kifunesina (columna 1 de la Figura 5.1.7) con una clara acentuación por parte R_{tm} , R_{twm} , R_{ms} y R_{msw} (columna 1, fila 3, 4, 5 y 6 de la Figura 5.1.7). Los cuatro factores agudizan la diferencia entre los dos mínimos descritos como ocurre con la energía, pero sin cambiar la localización en el eje de las abscisas.

Como ya se ha venido comentando para el caso del salacinol, la estructura determinada por DRX y el resto se diferencian lo suficiente como para considerar dos confórmeros. En el 2005, a partir de los mismos datos de STD-NMR y mediante el uso de otra metodología se llegó a un resultado similar (Wen 2005). La cristalografía y los datos de STD-NMR describen la cadena acíclica del salacinol con dos conformaciones desiguales. Aun así, los factores R-NOE describen un ligero mínimo entorno a la estructura cristalográfica, lo que nos induce a pensar que el confórmero descrito por DRX puede presentar una aportación menor a las intensidades del STD-NMR. Para poder verificar

esta hipótesis, habría que realizar un estudio más profundo del aporte de la saturación de ambos confórmeros a los datos de STD-NMR.

El caso de la epotilona-A es más complejo. Según los datos representados en la Figura 5.1.7, el mínimo global de los STD-NMR crudos cae en uno de los mínimos locales descrito por los factores R-NOE. Curiosamente, aunque los ángulos diedros parecen modificar significativamente las torsiones entre un mínimo y otro, los RMSD son pequeños. En un intento de explicar estos hechos, las estructuras obtenidas por todos los métodos fueron representadas en la Figura 5.1.8. En ella se observa claramente la existencia de dos conformaciones diferenciadas por la rotación de unos 113° del átomo H17 (Panel B). Éste cambio angular implica, a su vez, la rotación del metilo $\text{CH}_3\text{-27}$ y leves cambios estructurales en la cadena lateral. Pero este resultado no es nuevo; en el trabajo de Canales se llegaron a determinar dos confórmeros similares para la epotilona-A mediante datos de TR-NOESY (Canales, Nieto et al. 2014). Este resultado se confirmó en un reciente estudio donde se analizan nuevos derivados sintéticos de la familia de epotilonas (Jantsch, Nieto et al. 2016). En ambos estudios, se comentó que la epotilona-A unida a la β -tubulina es capaz de coexistir en dos conformaciones independientes, la sin- y la anti-. Además, se llega a determinar las proporciones entre confórmeros. Los resultados arrojaron que la conformación sin- es la más probable existiendo en una proporción del 84%. En nuestro trabajo, mediante el uso de los factores R-NOE, llegamos a identificar ambas conformaciones, aunque se diferencia la conformación anti- como más probable según la profundidad de los mínimos.

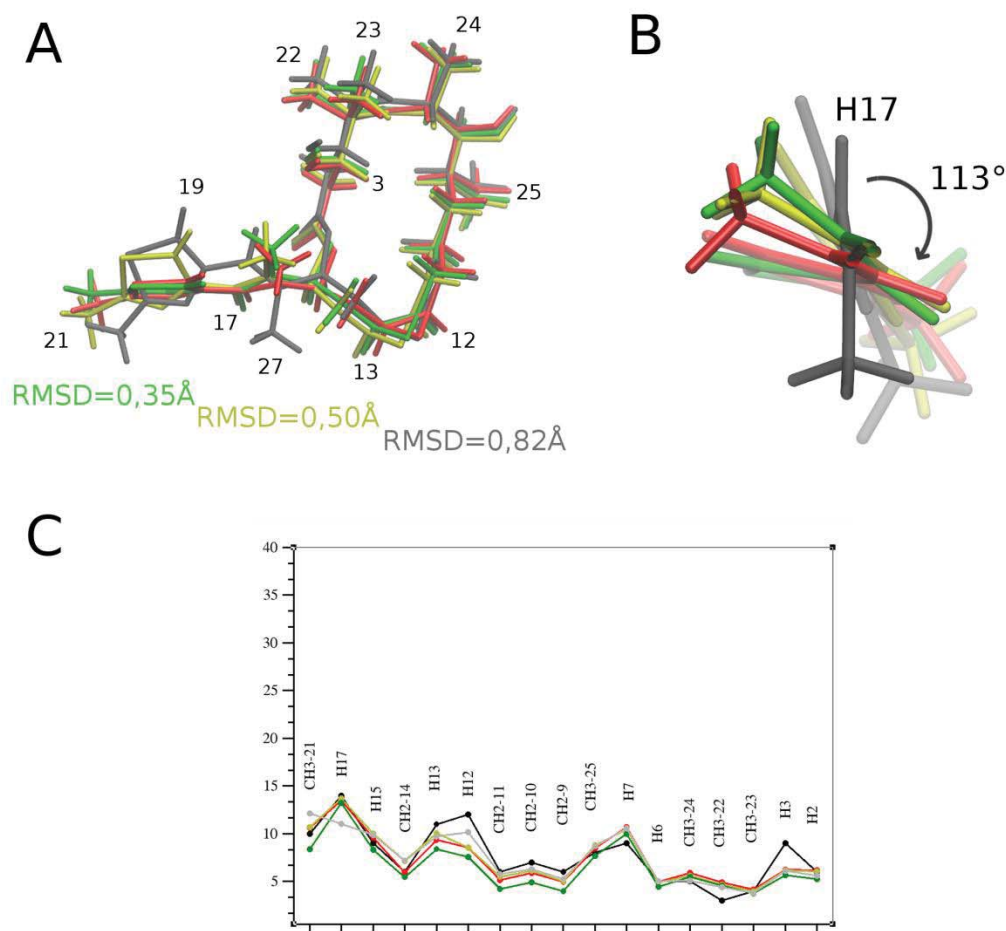


Figura 5.1.8. Representación de los conformeros del complejo de epotilona-A- β -tubulina obtenidos mediante el uso de técnicas de DRX (rojo), modelado molecular (verde), STD-NMR (amarillo), factores R-NOE (gris). (A) Solapamiento de las coordenadas de los conformeros de la epotilona-A (B) Ampliación de la zona transversal de la cadena lateral de la epotilona-A. (C) Datos de STD-NMR calculados a partir de las coordenadas de cada método de evaluación y los datos de STD-NMR experimentales (en negro). Los RMSD se obtienen tomando como referencia la estructura cristalográfica.

Los resultados obtenidos en este último estudio sorprenden pues, los factores R-NOE tuvieron la sensibilidad suficiente para poder diferenciar entre los conformeros descritos por TR-NOESY, la estructura cristalográfica y el campo

de fuerzas OPLS2005; por tanto, se demuestra que el STD-NMR no solo depende los datos de interacción intermolecular. Por eso y como se ha anunciado anteriormente, el análisis de los resultados de un experimento de STD-NMR tiene que ser cuantitativo. Además, señala que el acoplamiento entre los protones del ligando influye lo suficiente como para diferenciar conformeros con un RMSD de 0.8 Å; en este estudio, la cantidad de datos son suficientes para cubrir los grados de libertad conformacionales de todos los ligandos. Por tanto, las soluciones halladas por los factores R-NOE, a partir de la exploración del espacio conformacional, son completamente analíticas. A su vez, el estudio realizado con la epotilona-A complementa las conclusiones del trabajo realizado con TR-NOESY (Canales, Nieto et al. 2014).

5.1.7 Evaluación de los factores R-NOE en el reconocimiento de estructuras conocidas ligando-receptor

En la literatura se puede encontrar que los protocolos exhaustivos dedicados a la resolución estructural a partir de datos de STD-NMR siguen una serie de pasos que se inician con el acoplamiento de las moléculas, seguidos de la exploración conformacional del ligando en el sitio de unión y culminan con la relajación estructural. En este ámbito, las aplicaciones de los factores R-NOE en la evaluación estructural pueden ser diversas. Por tanto, el estudio del comportamiento de cada factor R-NOE, en cada paso, nos permitirá conocer el factor R-NOE apto para cada tarea.

Hasta ahora, el comportamiento de los factores R-NOE se estudió frente a la variación de los parámetros químico-físicos de la matriz CORCEMA. Además, se evaluó el desenvolvimiento de cada uno, en la exploración conformacional de tres ligandos en su sitio de unión. Los resultados de estos estudios demostraron que los factores actúan como funciones de pseudoenergía y son útiles tanto en la búsqueda de conformeros como en la optimización estructural

de complejos moleculares. Basados en estos resultados, nos planteamos evaluar los seis factores R-NOE en el procedimiento de acoplamiento molecular para sus posteriores usos como funciones de puntuación.

Este rol para el factor R_e está más extendido, dado que es muy común su uso en la evaluación de los acoplamientos moleculares de menor energía resueltas por algoritmos de modelado molecular (Barra, Jiménez et al. 2016, Marchetti, Perez et al. 2016). En estudios de esta índole, los parámetros químicos-físicos utilizados se obtienen por ajustes empíricos o ajustes entre rangos cortos de valores mediante la optimización del factor R_e .

La decisión de escoger cualquiera de las dos opciones se basa en la relación entre el tiempo de cómputo y la exactitud de los resultados. Como algunas de nuestras funciones mejoran su robustez frente a la variación de los parámetros en el CORCEMA; nos planteamos evaluar los seis factores R-NOE, en el procedimiento de búsqueda de la estructura descrita por datos de STD-NMR de los complejos proteína-ligando sin ajustes paramétricos previos. A consecuencia, los seis factores se sometieron a la búsqueda de las tres orientaciones en la unión de los complejos de kifunesina-dGMII, salacinol-dGMII y epotilona-A- β -tubulina entre 300 acoplamientos generadas con el AutoDock 2.4 (Hetenyi and van der Spoel 2002). En este estudio, se consideró un rango amplio de parámetros para resolver la matriz CORCEMA y las estructuras cristalográficas de los tres complejos proteína-ligando se sometieron a una optimización estructural. Las coordenadas resultantes fueron tomadas como referencia, de forma que por cada par de moléculas proteína-ligando se generaron 100 orientaciones. Para eliminar duplicados, los complejos obtenidos fueron agrupados por su RMSD, por lo que, finalmente, fueron seleccionadas 45 estructuras en la kifunesina-dGMII, 50 en el salacinol-dGMII y 62 en la epotilona-A- β -tubulina. Por cada estructura generada se realizaron 235 cálculos con el CORCEMA considerando diferentes tiempos de correlación rotacionales, constantes de velocidad de disociación y constantes de equilibrio. Los resultados fueron evaluados mediante la comparación de los

STD-NMR calculados y experimentales por medio de los seis factores R-NOE. Todos los valores obtenidos por cada factor se ordenaron en orden creciente. En cada caso, se considera que la estructura definida como “correcta” está ubicada en el primer puesto de la lista con el menor valor del factor R-NOE. Finalmente, se cuantificó el total de veces que se hallaron las estructuras de referencias en cada posición de las seis listas.

La mayoría de los estudios estructurales basado en los datos de STD-NMR no han sometido hasta la fecha, al factor R_e a condiciones tan desfavorables (una gran cantidad de estructuras y un amplio rango de parámetros). Bajo estas condiciones límite se obtiene la Figura 5.1.9. En ella se ilustra el número de veces que se encontraron las estructuras de referencia en una posición de la clasificación. Claramente, se puede observar en la imagen que ninguna de las funciones de puntuación fue capaz de escoger el 100% de las veces las estructuras referencia como primera opción. El mejor rendimiento se obtuvo con los factores R_e y R_t al encontrar solo el 50% de las estructuras de referencia.

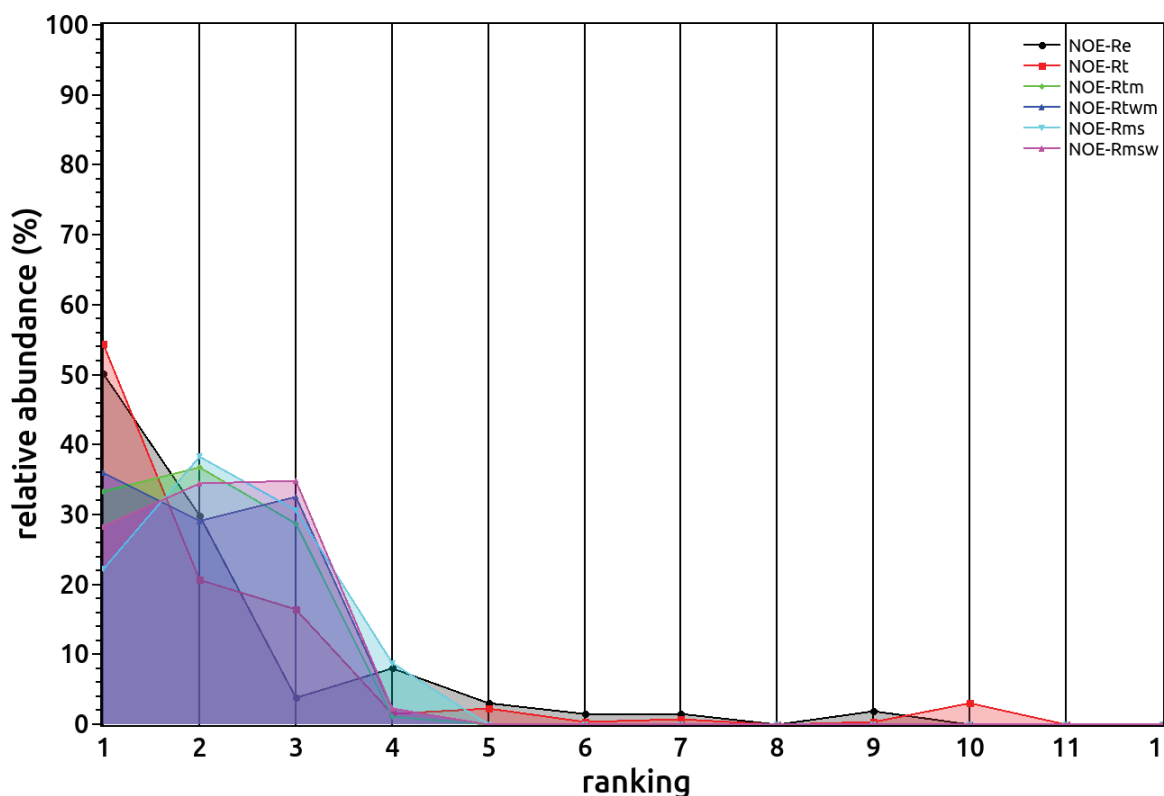


Figura 5.1.9. Capacidad de las funciones de ajuste de encontrar las estructuras de referencia para los complejos proteína-ligando entre varias poses de unión, considerando un amplio espectro de constantes químico-físicas. El eje de la abscisa corresponde a la posición de la lista. La posición 1 es la considerada como la estructura con mejor ajuste. El eje de las ordenadas es el número de veces que se encontró alguna estructura de referencia sin importar la parametrización usada.

Ahora, si en vez de tener en cuenta la primera posición, consideramos las cinco primeras posiciones de la clasificación. Encontramos que los factores R_{tm} , R_{twm} , R_{ms} y R_{msw} son capaces de localizar las estructuras de referencia el 100% de las veces. Este resultado implica que las cuatro funciones consiguen localizar las estructuras sin importar el conjunto de parámetros cinéticos o de relajación escogidos. En consecuencia, podríamos realizar una exploración exhaustiva con el ligando, sin necesidad de ajustes paramétricos, y sin reducir la exactitud

del método. Además, debido al número de orientaciones tenidas en cuenta y a la capacidad de rastrear correctamente el 100% de las estructuras frente al 95% alcanzado por el R_e , la implementación de estas funciones posibilitaría el aumento de la sensibilidad de los métodos de elucidación basados en datos de STD-NMR. Así, aseguramos que la estructura correcta siempre sea escogida cuando el algoritmo de acoplamiento molecular la genere.

Tabla 5.1.5. Porcentaje de acierto de las funciones de ajuste en el reconocimiento de estructuras de referencia entre 300 estructuras generadas por el AutoDock 4.2.

Posición	R_e	R_t	R_{tm}	R_{twm}	R_{ms}	R_{msw}
1	50±20%	54±5%	30±50%	40±40%	20±30%	30±30%
2	80.1±0.5%	80±10%	70±40%	70±40%	60±30%	60±30%
3	84±2%	92±4%	99±2%	98±3%	90±10%	98±3%
4	92±3%	93±5%	100%	100%	100%	100%
5	95±4%	95±4%	100%	100%	100%	100%

Un análisis más profundo y cuantitativo fue realizado sobre la cantidad de aciertos ubicados en las cinco primeras posiciones de cada lista. En el estudio se calculó el total de aciertos de los factores R-NOE en cada complejo molecular. Los resultados se promediaron y se calcularon las desviaciones estándar (Tabla 5.1.5).

Como ya hemos dicho, el primer puesto de cada lista tiene un valor bajo de aciertos, pero además se observa una elevada desviación entre los tres sistemas. Solo R_t muestra una precisión aceptable con un error relativo del 9%. Hay que recordar que la única diferencia entre R_e y R_t se encuentra en el término S_k . Por tanto, podemos adjudicar la mayor precisión de R_t a la

corrección del término S_k , quedando de manifiesto el papel de $S_{cal,k}$ en la corrección de las desviación por medio de la reducción de la incertidumbre en la formulación.

Pero, cuando se tienen en cuenta más posiciones, la capacidad de reconocer la estructura correctamente mejora en los tres sistemas estudiados y disminuye significativamente la desviación estándar entre ellos. A partir de la posición 3, los factores R_{tm} , R_{twm} , R_{ms} y R_{msw} logran incluir el 100% de las estructuras de referencia. Mientras que, R_e y R_t no son capaces de ubicarlas en las 5 primeras posiciones.

Sin embargo, las únicas funciones que mostraron ser significativamente menos sensibles a los parámetros del CORCEMA, en el estudio estadístico realizado en el apartado 5.1.5, fueron R_{tm} y R_{twm} ; en esa sección, también se pudo ver como R_{ms} y R_{msw} describen una tendencia a disminuir ligeramente sus variaciones máximas. Hay que recordar que el 50% de la contribución a R_{ms} y R_{msw} es aportada por un término que penaliza los ajustes paramétricos que no minimizan la diferencia media entre la saturación experimental y calculada. Esto se traduce en un aumento en el factor R-NOE cuando los parámetros se alejan de los valores reales. A consecuencia de la adición de este término, el efecto causado por ΔM o ΔM_w se ve reducido, y los dos factores no llegan a ser estadísticamente más robustos que R_e o R_t . A pesar de la contradicción interna en las formulaciones, los factores R_{ms} y R_{msw} alcanzaron los mismos resultados que R_{tm} y R_{twm} con solo el 50% de la aportación del término resta $(S_{exp,k} - (S_{cal,k} + \Delta M))^2$. Este resultado es un manifiesto de la sólida influencia ejercida por la modificación de este término. Además, constata que los términos ΔM y ΔM_w están directamente relacionados con la disminución de la sensibilidad de los factores R-NOE frente a los parámetros químico-físicos.

Como muestran los resultados, cuando no se conoce la parametrización, a partir de cualquiera de los factores R_{tm} , R_{twm} , R_{ms} y R_{msw} se puede obtener la estructura descrita por las intensidades de STD-NMR pues, las cuatro

funciones siempre localizan la estructura buscada entre sus primeras 5 opciones. A consecuencia, podemos plantear un protocolo de acoplamiento molecular en dos pasos. En el primer paso, se explora exhaustivamente la superficie de la proteína con un juego de parámetros escogidos empíricamente. Las cinco estructuras que minimizan más el valor de R_{tm} , R_{twm} , R_{ms} y R_{msw} son escogidas y se someten a un segundo paso, en donde se optimizan los parámetros cinéticos y de relajación con R_t o R_e . Ambas funciones han demostrado ser estadísticamente sensibles a los cambios en la parametrización, de forma que minimiza su valor en los casos que el ajuste disminuye las diferencias entre los STD-NMR experimentales y calculados. Finalmente, R_t y R_e escogerían la estructura que mejor describe las intensidades de STD-NMR, pues han recopilado el mayor índice de acierto entre sus dos primeras posiciones (~80%) (Tabla 5.1.5). Debido a que no conocemos el comportamiento de los factores inmersos en un procedimiento de ajuste paramétrico acoplado a una búsqueda estructural, en el próximo capítulo analizaremos el segundo paso propuesto. A causa de los resultados obtenidos en este trabajo, la implementación de los nuevos factores R-NOE en protocolos de acoplamiento molecular permitirá identificar con mayor certeza la estructura de los complejos moleculares descrita por los datos de STD-NMR.

5.2 Diseño e implementación del STD-MaPa, software para la determinación estructural de complejos moleculares con datos de RMN

5.2.1 Antecedentes

A día de hoy, no existe un algoritmo automático que sea capaz de crear y evaluar la estructura descrita por los datos de STD-NMR a partir de las coordenadas de macromoléculas y ligandos. Además, la ausencia de un protocolo autoconsistente encargado de esta tarea, aumenta la posibilidad de cometer errores. Como se puede extraer del estudio conformacional del apartado 5.1.6, la carencia de una correcta exploración del espacio de coordenadas puede inducirnos a escoger una estructura errónea. Por tanto, el diseño de una metodología que se pueda implementar en un algoritmo automático y autoconsistente es imprescindible, si se desean obtener resultados analíticos a partir de los espectros de STD-NMR.

Hasta el momento, hemos estudiado cinco nuevos factores R-NOE que mejoran la robustez, sensibilidad y exactitud del factor R-NOE tradicional. Además, analizamos sus comportamientos aplicándolos activamente en protocolos de exploración conformacional e identificación de la orientación del ligando en los procedimientos de acoplamiento molecular. Los avances logrados nos dan los soportes necesarios para el diseño de una rutina analítica y automática que identifique la estructura descrita por el ligando en la unión a partir de los datos de STD-NMR para complejos formados entre macromoléculas y ligandos.

Por estas razones, uno de los objetivos de nuestro trabajo es cubrir el diseño, implementación y comprobación de un algoritmo automático, capaz de obtener la estructura descrita por los datos de STD-NMR. El primer paso estuvo centrado en la creación de una metodología que se pudiera automatizar y que asegure la veracidad de los resultados. La implementación desencadenó en la

creación de un entorno único dedicado a la elucidación de las estructuras de diversos tipos de complejos moleculares basada en datos de STD-NMR. Posteriormente, se realizaron diversos protocolos enfocados en la comprobación del algoritmo diseñado.

5.2.2 Diseño de la metodología

Los protocolos de obtención estructural y su posterior evaluación a partir de datos de STD-NMR están muy estandarizados. A grandes rasgos, podemos decir que en ellos se realiza un acoplamiento molecular, seguido de la optimización estructural del complejo molecular y se termina con la comparación de los datos de STD-NMR experimentales y los calculados con el CORCEMA-ST (Canales, Rodríguez-Salarichs et al. 2011). Como ya se ha comentado, estos métodos utilizan pasivamente la información extraída de los experimentos de STD-NMR y carecen de los mecanismos para asegurar la veracidad de las estructuras.

Sin embargo, la implementación de dichos protocolos en un algoritmo automático es fácil. Además, la combinación de la automatización con la adición de ciertos controles permite, como se constatará más adelante, gozar de un algoritmo de elucidación estructural cuantitativo que mimetiza el uso activo de los datos de STD-NMR y asegura la veracidad de las estructuras. Por tanto, la metodología implementada en el presente trabajo aglutina las tareas comentadas anteriormente junto a otras que se fusionan en el protocolo esquematizado en el diagrama de flujo de la Figura 5.2.1.

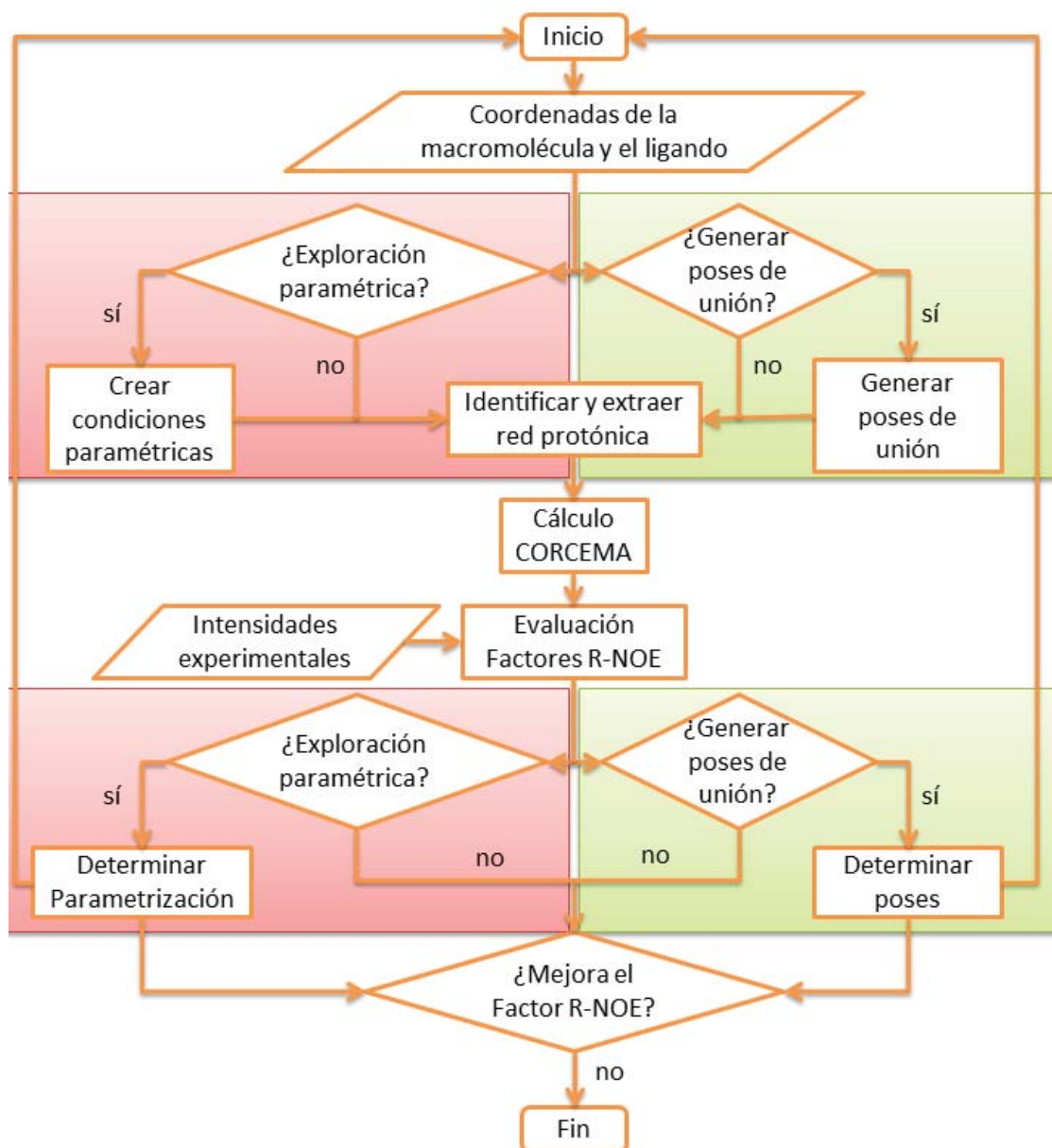


Figura 5.2.1. Diagrama de flujo del método implementado en el programa STD-MaPa (*STD Modelling Adjusting Parameter Approach*): Los óvalos representan el inicio y fin del programa, los romboides simbolizan cualquier tipo de entrada de datos, los rombos indican operaciones lógicas de respuestas binarias (sí o no) y los rectángulos simbolizan cualquier tipo de operación. Diferenciamos con fondo rosado, los elementos del método relacionados con la exploración paramétrica y con fondo verde, los elementos del método relacionados con la generación de las coordenadas del ligando en la unión.

El protocolo comienza obteniendo las coordenadas de las estructuras de la macromolécula y el ligando. Posteriormente, se aplican dos operaciones lógicas de respuesta binaria para decidir si se generaran diversos entornos paramétricos y diversas orientaciones. Una vez llegado a este punto, los protones de ambas moléculas son identificados según el grupo químico al que están enlazados, y se procede a extraer sus posiciones espaciales. Con estos valores, se calculan los datos de STD-NMR para las diferentes combinaciones protónicas mediante el método CORCEMA, tomando en consideración los entornos paramétricos generados. Los seis factores R-NOE estarían implementados para poder cubrir los diversos procedimientos incluidos en la metodología. De esta forma, obtenemos las intensidades experimentales y evaluamos los datos calculados con los factores R-NOE. Finalmente, la metodología arroja las coordenadas y los parámetros químico-físicos que, en su conjunto, minimizan los factores R-NOE. A partir de aquí, la metodología da diversos caminos que se pueden recorrer hasta alcanzar el punto en el cual los factores R-NOE no disminuyan más sus valores. Además, la flexibilidad inherente en esta metodología permite la inclusión de cálculos de Dinámica Molecular o de exploraciones conformacionales.

5.2.3 Componentes del software STD-MaPa

Para lograr tal objetivo, varias rutinas de Fortran, Scripting Bash, tcl/tk y Matlab/Octave fueron diseñadas e implementadas para poder cubrir las necesidades de un protocolo tan ambicioso. En su conjunto, los algoritmos se encargarán de realizar las tareas que se relatan a continuación. Primero, se deben cargar las coordenadas iniciales de la estructura de la proteína y el ligando y con ellas, crear un conjunto de modos de unión que asegure la generación de la estructura descrita por los datos de STD-NMR. Paralelamente, se generan diversos entornos paramétricos. A partir de las coordenadas del ligando en el sitio de unión, se recogen las de los protones de

las redes protónicas de las moléculas y se identifica su origen químico. Luego, se realiza el cálculo del STD-NMR mediante el método CORCEMA, combinando las redes protónicas y los entornos paramétricos generados, para que finalmente, con todos los datos de STD-NMR calculados, se evalúen con los seis factores R-NOE. En este paso, el protocolo tiene siempre que escoger la estructura descrita por los datos de STD-NMR experimentales. Además, la combinación de las rutinas tiene que aportarle al protocolo la suficiente flexibilidad para que se puedan acoplar algoritmos externos para el pos procesado.

Las rutinas que hemos diseñado no realizan esta tarea debido al cuantioso número de algoritmos existentes en la literatura, que pueden ser utilizados en la exploración y optimización estructural de las estructuras resueltas por nuestro protocolo. Además, como se verá en el apartado **5.3**, esta decisión permite acoplar nuestra metodología a diferentes métodos de resolución estructural de RMN, DRX o crio-ME.

Por último, el procedimiento metodológico ilustrado en el diagrama de flujo de la Figura 5.2.1 ha sido implementado en el programa, hecho por nosotros, STD-MaPa. El programa STD-MaPa obtiene de forma automática y bajo un entorno amistoso la estructura de los complejos macromolécula-ligando. Como filosofía, se persigue proporcionar una interfaz accesible para usuarios no expertos. Por consiguiente, la instalación de los paquetes es automática y está bien documentada en el fichero de texto "INSTALL". Como la instalación acarrea la compilación de diversos programas y se desea la máxima compatibilidad con todos los sistemas, la instalación se realiza mediante los simples comandos de consola "*configure*" y "*make*". Así, cualquier usuario con conocimientos básicos en la consola de Linux o Unix, va a poder instalar el programa. Para usuarios de PC, la versión 0.9 del programa se probó satisfactoriamente en Windows 7 mediante el Cygwin 2.5.2.

Las diferentes rutinas que forman parte del STD-MaPa se encuentran representadas esquemáticamente en la

Figura 5.2.2. En la ilustración se puede observar como el STD-MaPa está compuesto por disímiles módulos. Esta construcción está pensada para su fácil actualización, adaptación y modificación. Es necesario dejar constancia que el programa presenta rutinas desarrolladas por terceras personas que están bajo la Licencia Pública General (GNU) o en el caso que se aplique, se obtuvo el consentimiento expreso para su implementación. El uso de las rutinas creadas por terceras personas se ha debido al deseo de utilizar algoritmos eficientes y a la reducción del tiempo de programación. La licencia está bien documentada en el fichero “*LICENSE*” y viene junto con el instalador.

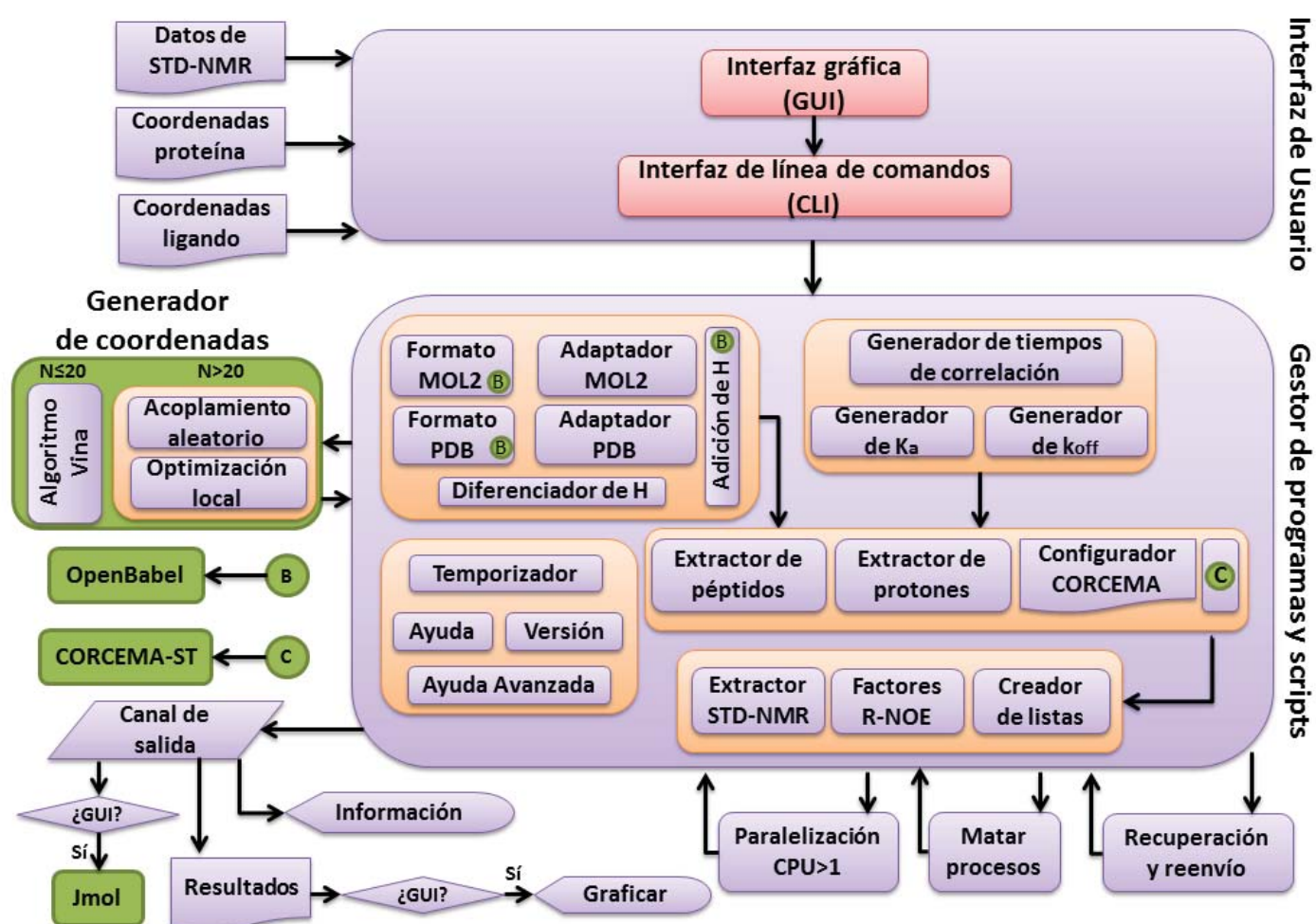


Figura 5.2.2. Constitución esquemática del programa STD-MaPa. En rosado, la interfaz accesible para el usuario; en naranja, módulos utilizadas por el gestor de programas; en morado, subrutinas; en verde, rutinas de terceros incluidas en el STD-MaPa. Las rutinas de terceros están bajo la Licencia Pública General (GNU) o el autor ha dado el consentimiento expreso para su implementación. STD-MaPa es un programa de uso libre y se distribuye bajo la licencia GNU.

El STD-MaPa se compone de diversos niveles de programación. El nivel superior está formado por la *Interfaz de Usuario*. A través de ella, los usuarios acceden al *Gestor de Programas y Scripts*. A su vez, la *Interfaz de Usuario* se compone de la *Interfaz Gráfica* (GUI) y la *Interfaz de Línea de Comandos* (CLI). La GUI está programada en Tcl/Tk y pretende ofrecer un entorno amigable que posibilita que los usuarios interactúen con las rutinas del programa cómodamente (ver Figura 5.2.3 y Figura 5.2.4-A). Está compuesta por funciones de tcl/tk, capaces de crear un entorno único que da la posibilidad de diseñar el proyecto, realizar el cálculo, recolectar los datos, graficarlos y analizar las estructuras de salida.

La Figura 5.2.3 muestra alguna de las facilidades que proporciona el STD-MaPa en la creación del proyecto. Es necesario conocer que la GUI no opera directamente sobre el *Gestor de Programas y Scripts*, solo envía las directrices a CLI. A su vez, CLI se compone de un único script programado en Bash encargado de ejecutar el *Gestor de Programas y Scripts*. Por tanto, se puede decir que el cálculo inicia cuando la CLI es ejecutada.

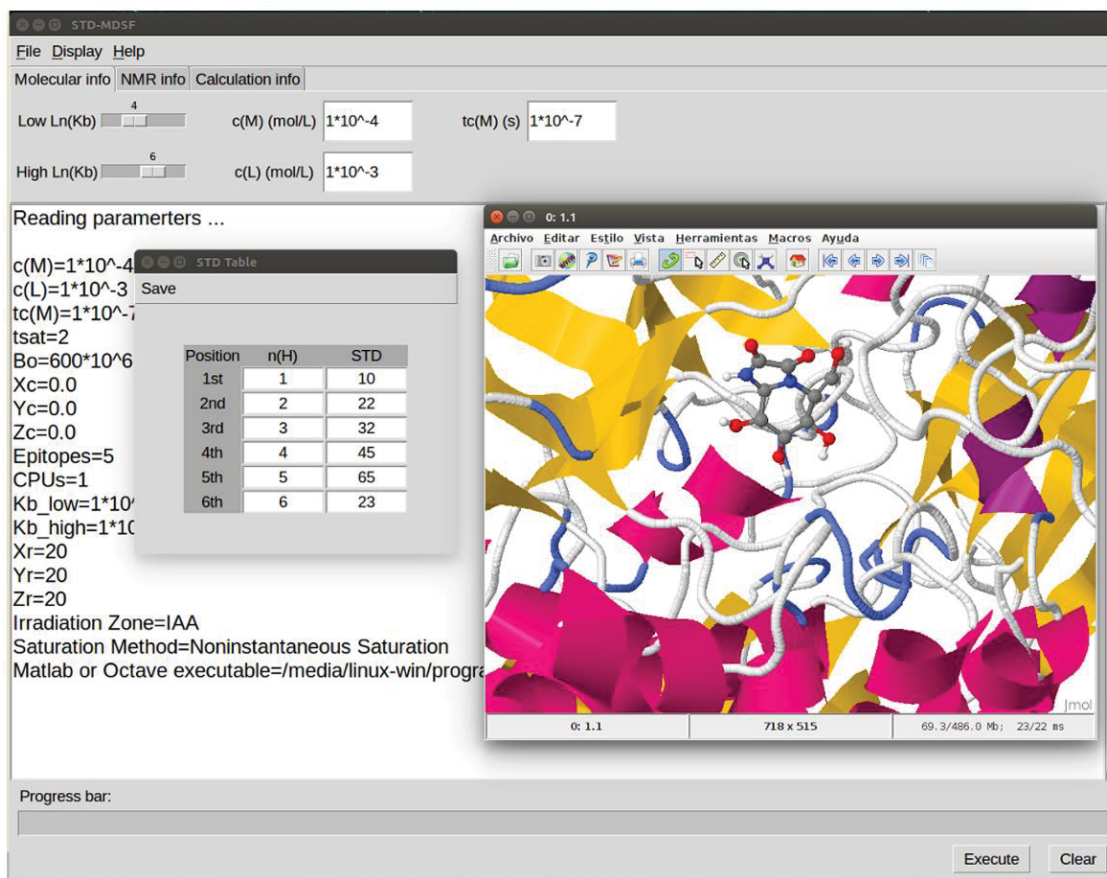


Figura 5.2.3. Interfaz gráfica del STD-MaPa, en la que se muestra la ventana *Tabla* y el *Visualizador de Moléculas* (Jmol).

Como se puede notar en la

Figura 5.2.2, en un nivel intermedio se encuentra el *Gestor de Programas y Scripts*. Este módulo es el centro del STD-MaPa y está encargado de gestionar la mayoría de los módulos del STD-MaPa. La ejecución de los programas Vina, OpenBabel (O'Boyle, Banck et al. 2011) y CORCEMA-ST también está ligada al gestor, aunque sean programas externos. Solo, las rutinas de *Paralelización*, *Parado de Procesos* y *Reenvío de Cálculos* se encuentran fuera de éste, debido a que ejercen un control directo sobre su trabajo.

El *Gestor de Programas y Scripts* regula 5 módulos internos y el *Generador de Coordenadas*. Entre ellos se encuentra un módulo utilitario que contiene el temporizador, las ayudas y la información sobre la versión del programa. El tratamiento del formato de los ficheros corre a cargo de un módulo que está interconectado con el OpenBabel. Este módulo es de gran importancia porque ejerce la función de intérprete entre la entrada y salida del resto de módulos. Aquí, el OpenBabel se utiliza para traducir los sistemas de coordenadas y, junto con el *Diferenciador de Hidrógenos*, rectifica los estados de protonación de las estructuras. Todos los programas y subrutinas de este módulo están escritos en Fortran; el resto de los módulos están ligados a la metodología de elucidación de estructuras y van a ser comentados con mayor profundidad en los siguientes apartados.

5.2.3.1 Módulo de optimización paramétrica

La optimización paramétrica se divide en dos procesos ubicados metodológicamente antes y después del cálculo de las intensidades teóricas. En la primera parte se construyen los diferentes entornos paramétricos que se desean explorar. El módulo está constituido por tres funciones que se encargan de generar los rangos de valores para el tiempo de correlación rotacional del ligando, la constante de asociación y la constante de velocidad de disociación. El tiempo de correlación de la macromolécula no se incluyó en la primera versión del programa porque puede obtenerse por medio del HYDROPRO (Garcia de la Torre, Navarro et al. 1994, Garcia de la Torre, Huertas et al. 2000). Aunque, como podrán ver posteriormente, podemos realizar su optimización a través de un script externo. La Figura 5.2.4 muestra las facilidades que nuestro entorno gráfico ofrece para configurar los rangos fácilmente.

Posterior al cálculo de las intensidades de STD-NMR, los datos pasan al módulo que contiene los factores R-NOE. Aquí se organizan los parámetros y las orientaciones obtenidas en orden creciente, a partir de los valores de los factores R-NOE. Por lo general, en las primeras posiciones se pueden encontrar todos los acoplamientos con un mismo juego de parámetros. Esta aplicación del STD-MaPa nos permite determinar de manera aproximada la constante de asociación aparente, la constante de velocidad de disociación aparente y los tiempos de correlación. En el apartado 5.2.4 se analiza la capacidad de optimización paramétrica del programa.

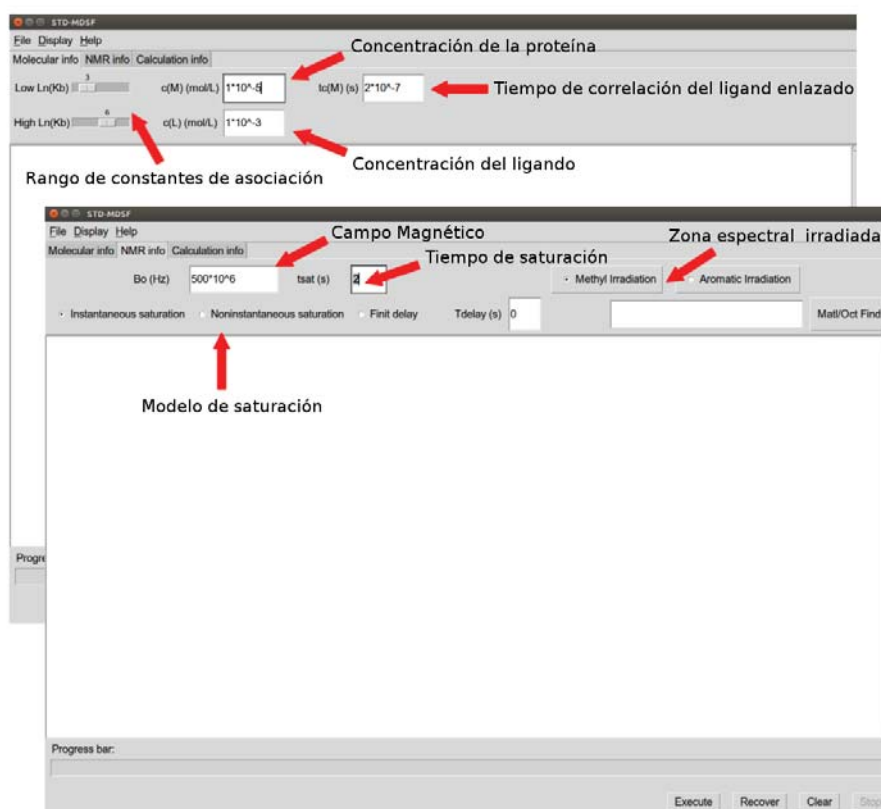


Figura 5.2.4. Interfaz gráfica del programa STD-MaPa dedicada al ajuste de las condiciones experimentales y parámetros químico-físicos. Se accede a ambas ventanas activando las pestañas de “*Molecular Info*” y “*NMR Info*”

5.2.3.2 Módulo de generación de modelos

La generación de la unión entre la macromolécula y el ligando es un paso crítico en nuestro protocolo y su diseño es complejo. El algoritmo involucrado muestra, en este paso, la creación de un complejo molecular estructuralmente cercano al descrito por los datos de STD-NMR en ausencia de un campo de fuerzas tradicional. Además, el algoritmo tiene que lograr su objetivo creando una mínima cantidad de complejos moleculares.

Si contamos con las rutinas que se encuentran en la literatura dedicadas a la generación de posiciones entre moléculas, el diseño de un nuevo algoritmo carece de sentido. Por lo tanto, optamos por la búsqueda de un algoritmo externo que cumpla con nuestras condiciones y sea fácilmente adaptable. A consecuencia, el AutoDock Vina fue escogido (Trott and Olson 2010).

El AutoDock Vina es un programa de código abierto dedicado al acoplamiento molecular y al cribado virtual. El programa fue escogido porque presenta una opción de exploración espacial aleatoria y otra de optimización estructural de búsqueda local (Baxter 1981, Blum, Roli et al. 2008). En la primera parte, el programa solo tiene en cuenta los choques estéricos y no considera ninguna función de energía. Mientras que, en su segunda opción, Vina realiza una minimización energética local optimizando los contactos entre la superficie de la macromolécula y el ligando. Además de estas dos secciones, el programa consta de un método principal que combina los anteriores pasos con optimizaciones estocásticas basadas en algoritmos genéticos (Holland 1975) y enfriamiento simulado ("*Simulated Annealing*") (Kirkpatrick, Gelatt et al. 1983). Por tanto, las funcionalidades del AutoDock Vina son idóneas para el diseño de un protocolo de acoplamiento molecular con los factores R-NOE como funciones de energía.

El método principal implementado en AutoDock Vina consta de una sucesión de pasos de mutaciones y optimizaciones locales, siendo cada paso aceptado acorde a los criterios de Metropolis (Hastings 1970). En su implementación se aplicó el método de optimización local Broyden-Fletcher-Goldfarb-Shanno (BFGS) (Wright and Nocedal 1999). El método BFGS es un eficiente método de cuasi-Newton que usa el valor de la función de puntuación y su gradiente para encontrar los mínimos locales. La combinación de todo esto da como resultado un algoritmo adaptado para una exploración rápida y eficiente de las superficies de las macromoléculas.

Estos dos mecanismos fueron implementados en nuestro programa pues, si solo se hubiese implementado el algoritmo de generación aleatoria, la exploración espacial consumiría mucho tiempo de cálculo o se tendría que considerar un espacio de exploración reducido. Ambas consecuencias son inaceptables debido a que reducirían la posibilidad de encontrar la estructura descrita por las intensidades de STD-NMR.

En conclusión, el STD-MaPa utiliza el método principal de Vina para genera las primeras 20 orientaciones de la unión. Así, aseguramos la descripción de las 20 posiciones menos energéticas en la zona explorada. Para el resto de las orientaciones, implementamos una personalizada sucesión de pasos de generación aleatoria de coordenadas, seguida de reducidas carreras de optimizaciones locales. En la generación de los acoplamientos, son aceptadas todas las estructuras que no presenten choques estéricos. A consecuencia de la adaptación de los algoritmos obtenidos por el AutoDock Vina, logramos un *Generador de Coordenadas* rápido y reproducible.

La exactitud del método de búsqueda de complejos moleculares del AutoDock Vina fue estudiada por sus autores y comparada con el AutoDock 4.0 (Morris, Goodsell et al. 1998). Aunque los resultados sorprenden por el sustancial aumento de exactitud, no son suficientemente buenos para nuestro algoritmo. En su estudio, solo el 78% de los 190 casos estudiados fueron predichos; en cambio, nuestro protocolo necesita un algoritmo que sea capaz de generar el

100% de las estructuras que se buscan. La causa de los fracasos pudo deberse a diversos factores, entre los que se pueden encontrar: el algoritmo de generación, el espacio conformacional explorado o la función de energía, entre otros. En cualquier caso, el procedimiento de nuestro protocolo para generar coordenadas tiene que ser estudiado para poder delimitar las condiciones que aseguran el 100% de las predicciones.

Por consiguiente, hemos probado la exactitud del *Generador de Coordenadas* con una base de datos de referencia de cribado virtual llamada Directorio de Señuelos Útiles o, con su nombre original, “*Directory of Useful Decoys*” (DUD) (Huang, Shoichet et al. 2006). DUD se diseñó para ayudar a comprobar los algoritmos de ensamblaje molecular mediante el uso de señuelos desafiantes. La base de datos contiene 2950 ligandos y 40 proteínas dianas. A su vez, cada ligando tiene 36 señuelos físicamente similares con topologías distintas, resultando en una base de datos compuesta de 98266 compuestos. Como se desprende de este estudio, la validez de nuestro programa solo se probará en sistemas compuestos por proteínas y ligandos.

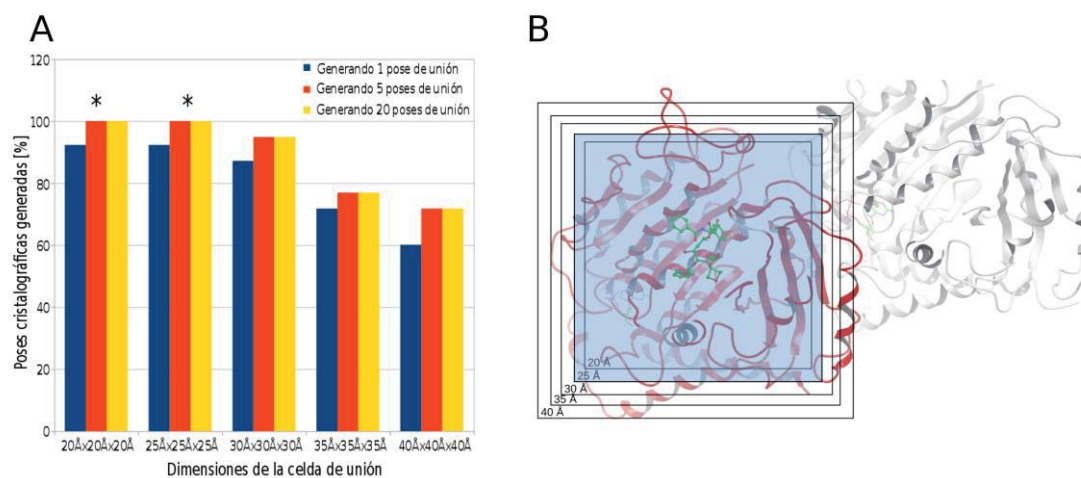


Figura 5.2.5. (A) Estructuras cristalográficas generadas con el *Generador de Coordenadas* a partir de los compuestos de la base de datos DUD. Cada complejo es contado como una respuesta positiva si al menos se llega a generar una vez la estructura cristalográfica. Las respuestas afirmativas se cuantificaron y normalizaron con respecto al total de complejos. (B) Representación esquemática de las dimensiones de los paralelepípedos en la β -tubulina (rojo). En azul, las dimensiones del espacio con el 100% de las estructuras cristalográficas generadas.

En nuestro estudio se ha tenido en cuenta todos los compuestos de la base de datos DUD. Los sitios de unión de los complejos moleculares fueron identificados y 5 paralelepípedos cúbicos de diferentes tamaños se centraron en los sitios de unión de cada diana. Las dimensiones de las aristas de las celdas se fijaron en 20, 25, 30, 35 y 40 Å. Luego, se realizaron tres experimentos con cada condición. Para ello, se le asignó al *Generador de Coordenadas* la tarea de crear 1, 5 y 20 modos de unión respectivamente. Finalmente, las estructuras de los complejos moleculares generadas correctamente fueron identificadas y con ellas, se cuantificaron la proporción de aciertos (Figura 5.2.5-A).

La Figura 5.2.5 muestra el porcentaje de aciertos por cada celda y número de orientaciones generadas. Con estos resultados se comprueba que el *Generador de Coordenadas* no es capaz de asegurar el 100% del éxito con todas las condiciones. Cuando se genera una única unión, el algoritmo sólo llega a generar el 90% de las estructuras cristalográficas en los paralelepípedos menores, debido a que, en las celdas de dimensiones mayores, el total de aciertos disminuye en función del tamaño. Estos resultados se asemejan al 78% obtenido por los autores del AutoDock Vina. La diferencia entre ambos estudios radica en el espacio explorado, puesto que en su estudio, las aristas eran siempre asimétricas y mayores a 22,5 Å. Sin embargo, en el nuestro, los aciertos menores al 80% se obtienen en celdas de aristas mayores a 30 Å, siendo, el mayor número de aciertos de nuestro estudio provocado por el uso de celdas cúbicas centradas en el sitio de unión. Esta simple consideración aumentó la eficiencia del algoritmo de acoplamiento molecular un ~10%.

Al generar 5 y 20 orientaciones alcanzamos a generar el 100% de las estructuras cristalográficas en las celdas de menor tamaño. Al igual que ocurre con la generación de una sola unión, en las celdas de aristas mayores a 25 Å disminuye la eficiencia. Por esa razón, en la celda con aristas de 30 Å, el *Generador de Coordenadas* logra crear solo el ~95% de las coordenadas cristalográficas. Esta limitación no es muy grave porque una celda con aristas de 25 Å proporciona el espacio suficiente para que el ligando rote libremente (Bursulaya, Totrov et al. 2003) y pueda explorar la cavidad de unión de la proteína. Con el fin de ilustrar las dimensiones de las cinco celdas, en la Figura 5.2.5-B se representa la superposición teórica de los paralelepípedos utilizados en nuestro estudio sobre el sitio de la β -tubulina. Como podemos confirmar, la celda con aristas de 25 Å permite explorar una extensa zona de la superficie de esta proteína.

5.2.4 Complejos moleculares de referencia

Posterior al análisis de *Generador de Coordenadas*, procedemos a evaluar la capacidad del programa en la optimización de los parámetros y en el reconocimiento de los modos de unión más cercanas a las descritas por los datos de STD-NMR. En estos estudios, al igual que en el apartado 5.1, se necesitan complejos moleculares para utilizarlos como referencia. Al igual que en ese apartado, los complejos moleculares tienen que cumplir con que los datos de STD-NMR y las estructuras atómicas (a alta resolución) hayan sido obtenidas previamente. Además, ambas técnicas tienen que describir la misma estructura atómica. Hasta ahora habíamos utilizados con éxito la kifunesina-dGMII, salacinol-dGMII y la epotilona-A- β -tubulina. Para este nuevo estudio, hemos seleccionado adicionalmente los complejos de podofilotoxina- α,β -tubulina y PM060327- α,β -tubulina. En conjunto, los cinco compuestos aportan una riqueza estructural que los convierte en un buen grupo de referencia (Figura 5.2.6).

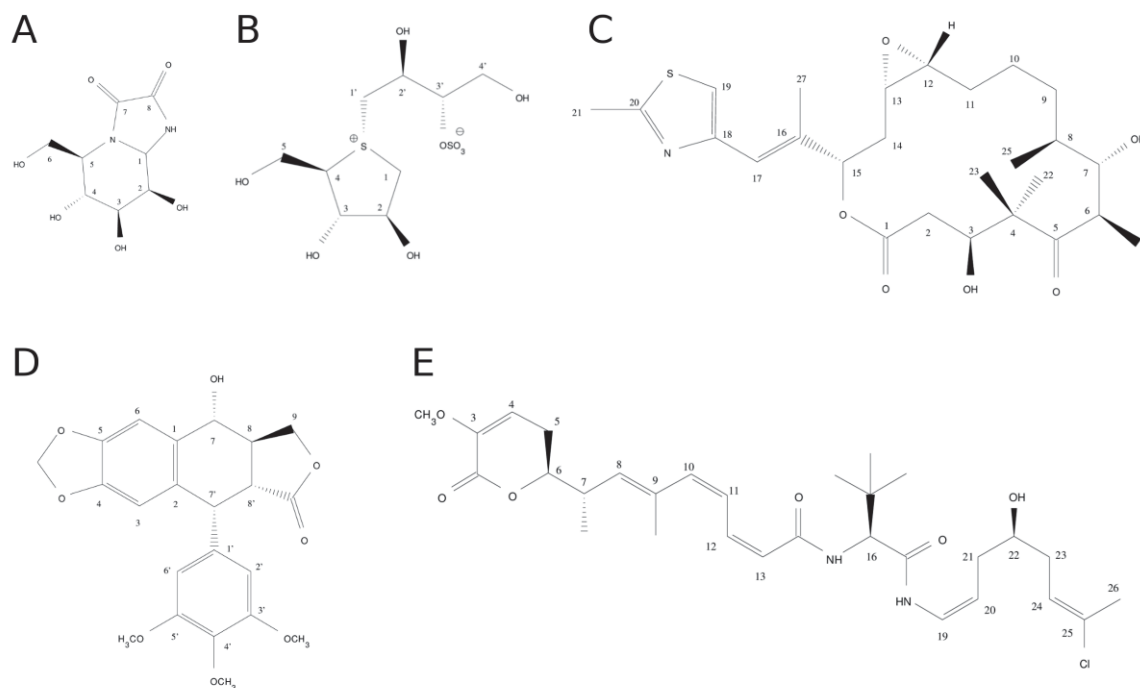


Figura 5.2.6. Estructura química de los ligandos de referencia. (A) Kifunesina, (B) Salacinol, (C) Eptilona-A, (D) Podofilotoxina y (E) PM060327.

Los nuevos candidatos son compuestos orgánicos Inhibidores del Ensamblaje de Microtúbulos que se unen a la interface entre α - y β - tubulina. La podofilotoxina (Figura 5.2.6-D) es un lignano natural muy conocido por su actividad antitumoral y antiviral. Este lignano se enlaza al sitio ubicado en la interface intradimérica del heterodímero de α,β -tubulina (Ravelli, Gigant et al. 2004). Mientras que, el PM060327 (Figura 5.2.6-E) pertenece a una reciente familia de compuestos citotóxicos de origen marino. Según el cristal del análogo PM060184 (Prota, Bargsten et al. 2014), esta familia se une a la superficie de β -tubulina que forma parte de la interface dimérica al constituirse los microtúbulos.

La estructura cristalográfica de la podofilotoxina fue resuelta en el 2004. Al igual que la eptilona-A, la podofilotoxina fue cristalizada junto con el complejo

T₂R. La resolución de la estructura cristalográfica es de 4.2 Å y su código PDB es 1SA1. En cuanto al PM060327, su estructura cristalográfica no se ha resuelto hasta la fecha.

Tabla 5.2.1. Tabla de datos cristalográficos del complejo PM060327-T₂R.

	PM060327
Data collection	
Space group	P212121
Cell dimensions	
<i>a</i>, <i>b</i>, <i>c</i> (Å)	103.3, 156.3, 180.3
α, β, γ (°)	90, 90, 90
Resolution (Å)*	49.5-2.8 (2.87-2.80)
No. reflections	73295 (5276)
<i>R</i>_{meas} (%)	22.1 (417.0)
<i>R</i>_{pim}	6.5 (181.6)
<i>I</i> / σ<i>I</i>	13.7 (1.0)
Completeness (%)	100.0 (100.0)
Redundancy	13.1 (13.6)
Refinement	
Resolution (Å)	47.8-2.8
No. unique reflections	73295
<i>R</i>_{work} / <i>R</i>_{free} (%)	17.3 / 22.8
Average B-factors (Å²)	
Complex	83.7
Solvent	65.8
Wilson B-factor	73.7
Root mean square deviation from ideality	
Bond length (Å)	0.009
Bond angles (°)	1.079
Ramachandran statistics^c	
Favored regions (%)	96.8
Allowed regions (%)	3.05

Outliers (%)	0.15
--------------	------

*Highest-resolution shell is shown in parentheses. °As defined by MolProbity (Davis, Murray et al. 2004)

En este trabajo, con la colaboración del grupo del Profesor Steinmetz se logró cristalizar la estructura del complejo de PM060327-T₂R. La cristalización se realizó mediante la aproximación experimental utilizada para la epotilona-A y la podofilotoxina. El sitio descrito por la estructura cristalográfica del PM060327 es el mismo que el análogo previamente cristalizado. Ambos se encuentran extendidos sobre la superficie interdimerica de la β -tubulina, asumiendo una conformación lineal y con el grupo terc-butilo de la posición 16 marcadamente expuesto hacia el solvente. La unión cristalográfica revela tres puntos de anclaje con la proteína, los cuales incluyen un puente de hidrógeno entre el grupo carbonilo en la posición 2 y los residuos Asn102 y Lys105 de la β -tubulina; un puente de hidrógeno entre el grupo carbonil ubicado en la posición 14 y la Val181 de la β -tubulina; y una interacción hidrofóbica entre el grupo metilo en la posición 7 del ligando y el bolsillo hidrofóbico formado por los residuos Asn101, Asn102, Val182, Phe404 y Tyr408 de la β -tubulina.

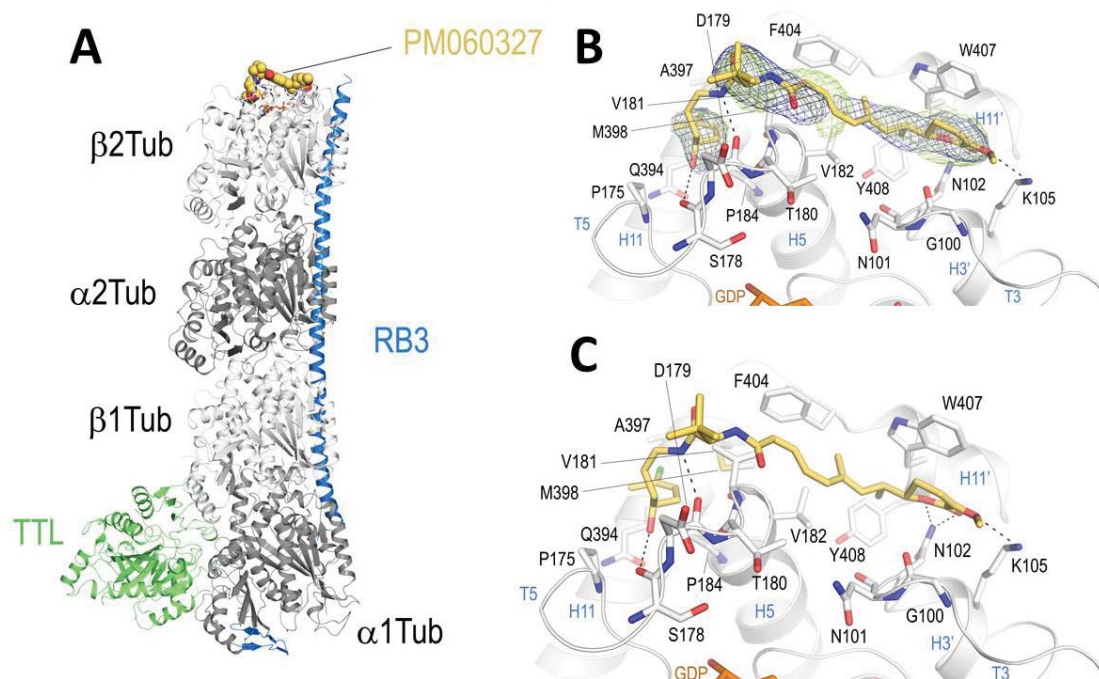


Figura 5.2.7. (A) Estructura cristalográfica del complejo de PM060327-T₂R. (B) Representación del sitio de unión con el modelo de la molécula de PM060327 localizada en la densidad electrónica. (C) Representación del sitio de unión con la estructura de la molécula de PM060327 unida a β2-tubulina.

La obtención y caracterización del STD-NMR de la kifunesina, salacinol y epotilona-A se han descrito a lo largo de este trabajo. Por tanto, solo procederemos a describir los complejos moleculares con la podofilotoxina y PM060327. Los valores de STD-NMR de ambos compuestos se representaron en la Figura 5.2.8 junto con los datos de la kifunesina, salacinol y epotilona-A.

Los espectros de STD-NMR de los complejos de podofilotoxina y PM060327 se adquirieron en presencia de heterodímeros de α,β-tubulina bajo condiciones despolimerizantes, utilizando un espectrómetro Bruker AVANCE de 600 MHz a 298 K y 310 K respectivamente. La unión de la podofilotoxina y de cuatro análogos naturales extraídos de la planta medicinal Mexicana *Bursera*

fagaroides fueron caracterizadas mediante datos de STD-NMR y TR-NOESY, en colaboración con el grupo de la Profesora Laura Álvarez de la Universidad Autónoma de Morelos (Antúnez-Mojica, Rodríguez-Salarichs et al. 2016).

Previo a la publicación de la estructura cristalográfica del PM060184, los mecanismos de acción y la ubicación de los sitios de unión de PM060184, PM050489 y PM060327 fueron investigados en un trabajo realizado en nuestro laboratorio. En éste, se caracterizó la unión del PM060327 mediante datos de STD-NMR y TR-NOESY. En el análisis realizado a los picos de cruces negativos del TR-NOESY, se detectaron correlaciones entre protones vecinos, no encontrándose correlaciones entre protones lejanos según la secuencia de la molécula. Esta evidencia, sugirió por primera vez la adopción de una conformación extendida. Resultado que posteriormente confirmó la estructura cristalográfica del PM060184.

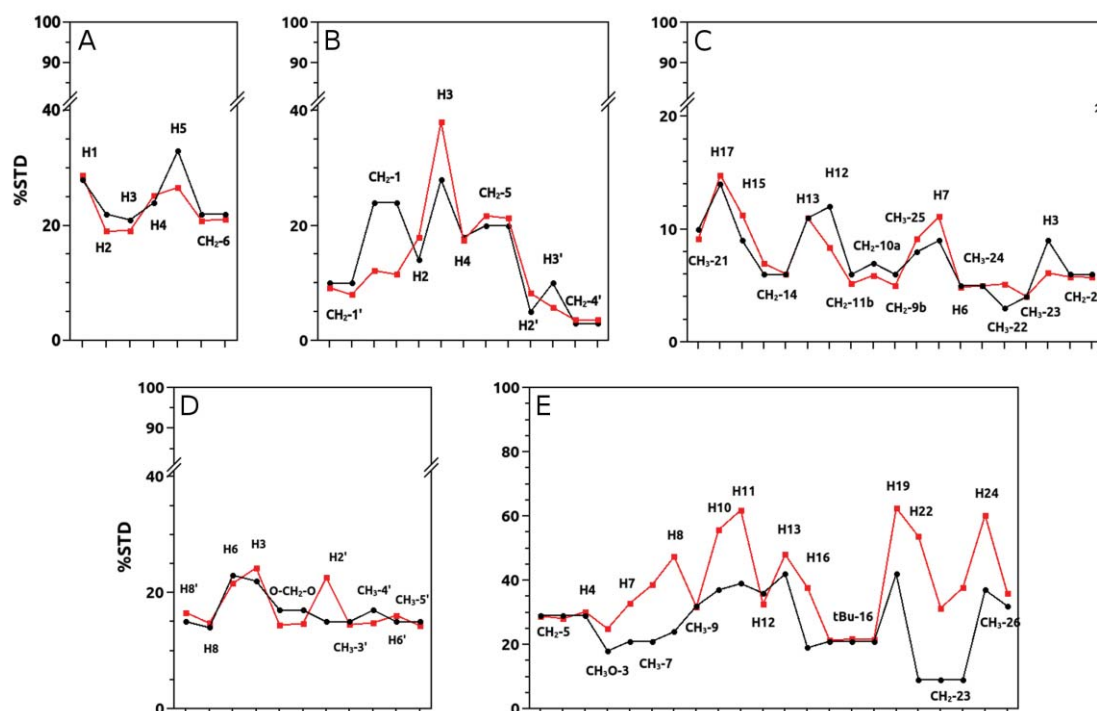


Figura 5.2.8. Datos de STD-NMR experimentales y calculados para los complejos de kifuensina-dGMII, salacinol-dGMII, epotilona-A-T₂R, podofilotoxina-T₂R y PM060327-T₂R. Los datos de STD-NMR experimentales están representados con líneas y círculos negros. Los datos de STD-NMR calculados están representados con líneas y cuadrados rojos.

Los datos de STD-NMR teóricos fueron calculados para la estructura optimizada de cada complejo molecular a partir de sus coordenadas cristalográficas. Como la estructura cristalográfica del salacinol y los datos de STD-NMR describen confórmeros diferentes; se eligieron las coordenadas referidas al mínimo global de la Figura 5.1.7.

Cuando comparamos los datos de STD-NMR experimentales y calculados en el complejo de PM060327 hayamos diversas inconsistencias. El sitio de unión expuesto al medio acuoso, la conformación lineal y la flexibilidad de algunas zonas de la estructura alejan las intensidades de STD-NMR calculadas de los datos experimentales. Como se puede ver, las intensidades de los protones

localizados en las regiones de mayor flexibilidad son sobreestimadas en nuestros cálculos. Puesto que el análisis que se va a realizar no tiene en cuenta la contribución de múltiples conformaciones, sólo son considerados los protones cercanos a los grupos que permiten el anclaje de PM060327 a la proteína. Por la importancia de las interacciones en estas regiones, la conformación de la estructura entorno a estos puntos estará más conservada. Al seleccionar los protones, también hay que considerar su posición en el ligando, ya que debemos asegurar la correcta descripción de la unión por medio de los datos de STD-NMR. Consecuentemente, se eligieron los protones H4, CH₃-O, H7, CH₃-9, H13, tBu-16 y CH₃-26.

Posterior al análisis de los datos de STD-NMR y la estructura cristalográfica de los cinco complejos moleculares, podemos continuar con las comprobaciones sobre los algoritmos implementados en el STD-MaPa. Por tanto, en el siguiente apartado estudiamos el comportamiento de los Factores R-NOE. Aunque los factores se han estudiado en profundidad, todavía no se ha analizado el desenvolvimiento de las seis funciones en un procedimiento de ajuste paramétrico ensamblado a uno de acoplamiento molecular.

5.2.5 Evaluación de la calidad de los factores R-NOE en el ajuste combinado de parámetros y estructuras de los complejos

En apartados anteriores se ha analizado el algoritmo implementado en el STD-MaPa dedicado al acoplamiento molecular. Con el análisis, se han conocido las condiciones indispensables para asegurar que el STD-MaPa genere la estructura descrita por los datos de STD-NMR. Pero, el programa tiene que generar al menos cinco orientaciones del ligando en el sitio de unión para reproducir la estructura buscada. Por tanto, el STD-MaPa debe ser capaz de identificar la orientación en el sitio que mejor representa los datos de STD-NMR. Debido los resultados de las pruebas anteriores, la implementación de

los seis factores R-NOE es la opción más adecuada. Por consiguiente, las seis funciones estarán a cargo de evaluar la calidad de las coordenadas y decidir la estructura que mejor describe las intensidades experimentales.

Hasta ahora, el desempeño de los factores R-NOE ha sido analizado en la evaluación estructural de complejos moleculares y los resultados desprendidos de estos estudios han sido muy buenos. Particularmente, ha sido llamativo el desempeño exhibido por los factores R_{tm} , R_{twm} , R_{ms} y R_{msw} . Pues, las cuatro funciones consiguieron identificar el 100% de las estructuras buscadas, sin importar el conjunto de parámetros cinéticos o de relajación escogidos. Partiendo de estos resultados, con la implementación de estos cuatro factores, el STD-MaPa podría ser utilizado en la exploración exhaustiva del espacio 3D entorno al sitio de unión, sin ajustes paramétricos previos, y sin reducir la exactitud del cálculo. A causa de las modificaciones matemáticas añadidas a los factores R-NOE (ΔM , ΔM_w , $N_i \Delta M^2$ y $N_i \Delta M_w^2$), éstas aseguran la presencia de la estructura descrita por el STD-NMR entre las cinco opciones con menor puntuación. Aunque es un buen resultado, estamos obligados a realizar una segunda identificación entre los cinco modos de unión para diferenciar cuál describe mejor los datos de STD-NMR. Por esto, estudiaremos otras alternativas metodológicas.

El desempeño del algoritmo encargado de evaluar las coordenadas del complejo molecular se estudiará con un amplio grupo de estructuras. En este paso, se pretende comprobar si el STD-MaPa logra optimizar los parámetros del CORCEMA, y al mismo tiempo, escoge correctamente las estructuras. Para poner a prueba esta sección del programa, la identificación de las estructuras de referencia de los cinco complejos se realizará a partir de 100 estructuras diseminadas en celdas cúbicas con aristas de 25 Å y 47 Å. Adicionalmente, el desempeño del algoritmo ha sido evaluado bajo dos condiciones paramétricas diferentes; en la primera y menos compleja se tomaron los valores de los parámetros que minimizan las diferencias entre las intensidades experimentales y las calculadas a partir de las estructuras de referencia,

mientras que, en el segundo caso, el algoritmo se tiene que encargar de optimizar los parámetros y al mismo tiempo identificar correctamente las estructuras de referencia.

La celda más pequeña tiene las dimensiones máximas en las que se pudo recrear el 100% de las poses cristalográficas de los complejos de la base de datos DUD. Por lo tanto, los resultados en esta celda son de vital importancia para nuestra metodología. Aunque conocemos que nuestro *Generador de Coordenadas* no asegura la correcta creación de las estructuras en una celda con arista de 47 Å. Esta celda ofrece la oportunidad de explorar una mayor expansión superficial en las proteínas. Así, el programa probaría su capacidad de elucidación entre diversos entornos de una misma proteína.

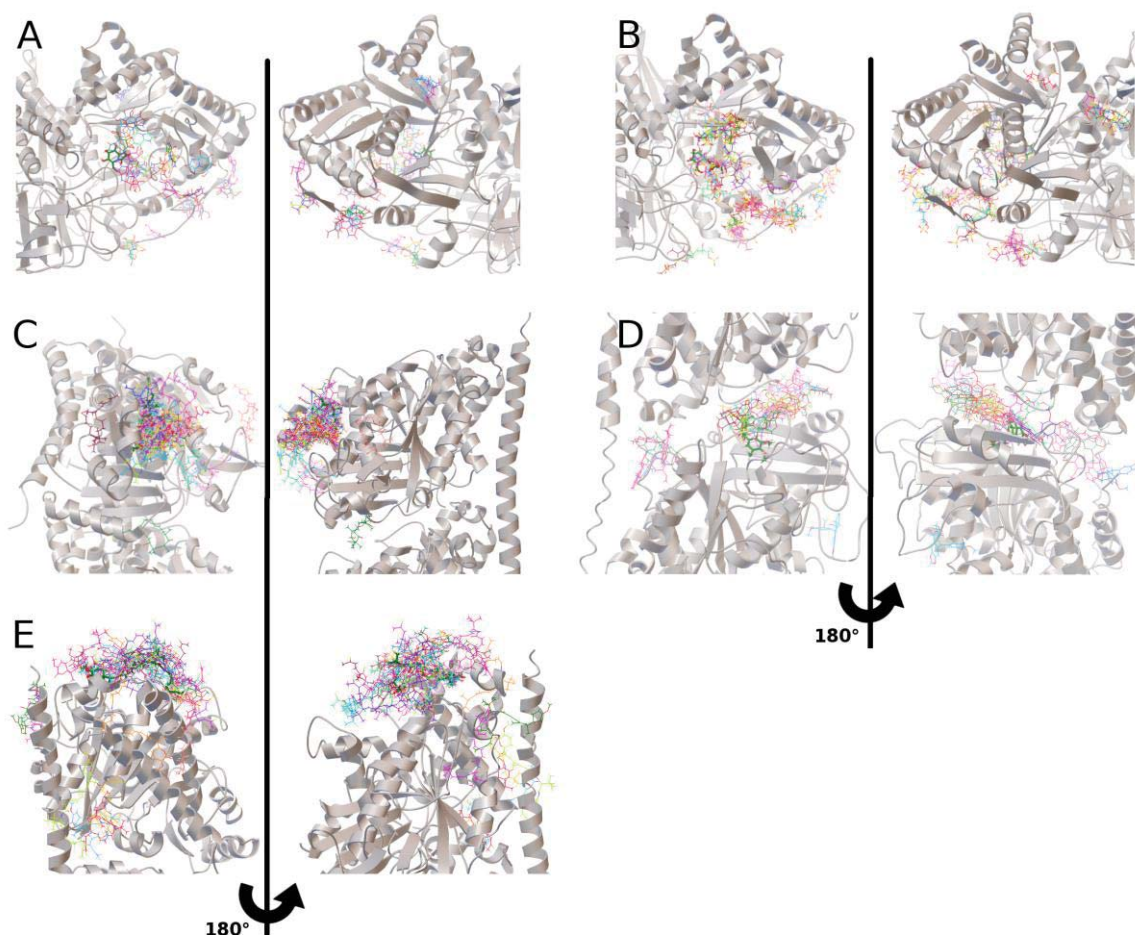


Figura 5.2.9. Superposición de las poses generadas por el Autodock 4.2 en dos cajas cúbicas centradas en el sitio de unión y de aristas igual a 25 Å y 47 Å. (A) Kifunesina enlazada a dGMII. (B) Salacinol enlazado a dGMII. (C) Epotilona-A enlazada a la β -tubulina. (D) Podofilotoxina enlazada a la interfaz intradimérica entre α - y β -tubulina. (E) PM060327 enlazada a la β -tubulina.

Para evaluar el rendimiento del algoritmo de evaluación, de manera independiente del paso de generación de coordenadas, escogemos la trayectoria sombreada en rosado perteneciente al mecanismo representado en la Figura 5.2.1. Debido a esto, los acoplamientos entre los ligandos y las proteínas de referencia se han tenido que generar externamente con el

AutoDock 4.2 en celdas cúbicas con aristas de 25 Å y 47 Å. La Figura 5.2.9 muestra, a grandes rasgos, la superposición de los modos de unión de ambas celdas en los cinco complejos moleculares. En la imagen se visualiza como los ligandos forman numerosas agrupaciones. Las agrupaciones se deben a la combinación del algoritmo genético Lamarckiano (Morris, Goodsell et al. 1998) para la búsqueda de estructuras de mínima de energía y a su campo de energía libre de origen semiempírico basado en un modelo termodinámico completo (Huey, Morris et al. 2007). La combinación favorece la aparición de las poses en las cavidades de las proteínas frente a otras zonas que se encuentran más expuestas al solvente, como se puede visualizar en la Figura 5.2.9. Este comportamiento mimetizaría el procedimiento de un protocolo real para la búsqueda extensiva entre diversas cavidades de la proteína. Además, asegurar que los ligandos estén situados en las cavidades de las proteínas, aumenta la similitud entre los entornos protónicos de cada pose, lo que ocasiona que el proceso de identificación sea más difícil.

La optimización paramétrica fue realizado para el tiempo de correlación del ligando libre, el tiempo de correlación de la proteína, la constante de asociación y la constante de velocidad de disociación (Tabla 5.2.2). El STD-MaPa siempre ajusta los tiempos de correlación del ligando libre y la k_{off} en función del rango de la constante de asociación definida por el usuario y el límite de la difusión en un líquido. La última versión del programa considera moléculas pequeñas y explora un rango de τ_c^l de 0,25 ns a 0,75 ns. El tiempo de correlación de la proteína se optimizó a través de un script externo que permitió que el STD-MaPa incluyera el tiempo de correlación de la proteína en el paso de optimización. Para el resto de las constantes ha sido escogido el rango máximo que permite el STD-MaPa.

Tabla 5.2.2. Rangos paramétricos explorados en los cinco sistemas de referencia

	Kifunesina- dGMII	Salacinol- dGMII	Epotilona A- β tubulina	Podofilotoxina- α,β tubulina	PM060327- β tubulina
$K_b [M^{-1}]^\dagger$	10^2-10^8	10^2-10^8	10^2-10^8	10^2-10^8	10^2-10^8
$k_{off} [s^{-1}]^\ddagger$	10^1-10^5	10^1-10^5	10^1-10^5	10^1-10^5	10^1-10^5
$\tau_{c-bound} [ns]^{**}$	100-900	100-900	10-90	10-90	10-9000
$\tau_{c-free} [ns]^\ddagger$	0.25-0.75	0.25-0.75	0.25-0.75	0.25-0.75	0.25-0.75

† Rango definido por el usuario

‡ Rango generado automáticamente por el programa

** Rango explorado mediante un script capaz de gestionar múltiples cálculos del programa

Los resultados fueron separados por factor R-NOE, tamaño de celda y las dos condiciones paramétricas asumidas. La relación entre las estructuras elegidas correctamente y el total generadas para cada complejo molecular fue cuantificada. La Figura 5.2.10 ilustra el porcentaje de aciertos alcanzados por los factores R-NOE en las cinco primeras posiciones de las seis listas generadas por el STD-MaPa.

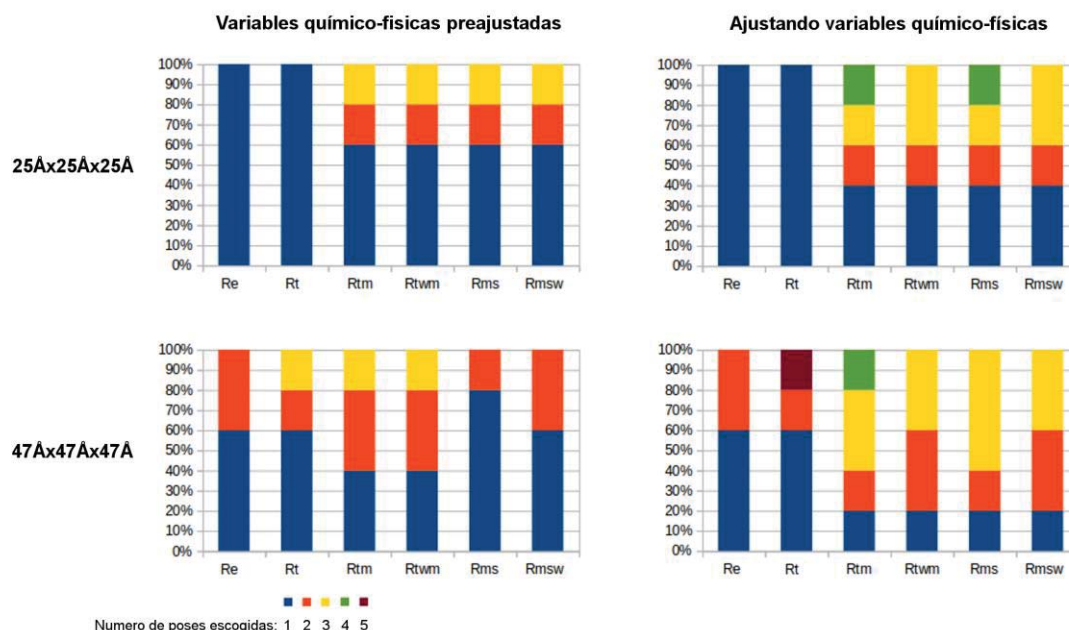


Figura 5.2.10. Relaciones entre las estructuras correctas y el total generadas para cada complejo molecular por cada Factor R-NOE. El estudio se ha realizado con cinco complejos proteína-ligando y dos celdas de aristas de 25 Å y 47 Å.

Los datos exponen una clara pérdida de exactitud por parte de todos los factores R-NOE cuando aumenta la superficie explorada. Por lo tanto, trabajar con celdas de menor tamaño va a beneficiar el funcionamiento del STD-MaPa, siempre que aseguremos la creación del suficiente espacio para que el ligando pueda explorar todo el sitio de unión y tenga libertad para rotar en la celda.

Ajustar los parámetros en la búsqueda de la estructura real no mejoró la sensibilidad de los factores R_{tm} , R_{twm} , R_{ms} y R_{msw} . Todos proporcionaron el 100% de las estructuras de referencia en las cuatro primeras posiciones de sus listas. Este resultado asegura que la sensibilidad de los cuatro factores no cambia con la modificación del método paramétrico a causa de su robustez.

En cambio, los factores R_e y R_t modifican su desempeño favorablemente con el ajuste paramétrico o con los parámetros preajustados. Las dos funciones encontraron el 100% de las estructuras de referencia en la celda más pequeña. En la celda mayor, el factor R_t acierta menos que R_e , dado que R_e aglutina todas las orientaciones correctas en las dos primeras posiciones de la lista.

Hasta ahora, hemos considerado que conocemos la localización del sitio de unión. Pero, cuando no se conoce, se necesita expandir la búsqueda a todas las cavidades de la proteína. Este cuadro es similar al descrito en la celda con aristas de 47 Å. En esta celda, fueron exploradas varias cavidades como muestra la Figura 5.2.9. Por tanto, si nos basamos en los resultados obtenidos con el factor R_e , el programa STD-MaPa está capacitado para evaluar los resultados de exploraciones más extensas.

En los procedimientos con parámetros químico-físicos prefijados, los factores aumentaron su sensibilidad. Indistintamente, R_e y R_t identificaron el 100% de las poses buscadas con o sin optimización paramétrica. Por tanto, se puede concluir que la metodología implementada en el STD-MaPa es capaz de identificar la estructura descrita por los datos de STD-NMR por ambas metodologías. Además, tiene la capacidad de optimizar los tiempos de correlación, la constante de asociación y la constante de velocidad de disociación a partir de las intensidades de STD-NMR experimentales.

5.2.6 Resolución de la estructura de los complejos problemas mediante el programa STD-MaPa

En esta sección, vamos a aplicar toda la información obtenida hasta el momento para resolver la estructura de complejos proteína-ligando con nuestro programa. Así podemos tener un programa que basado en datos de STD-NMR asegure la resolución de la estructura de complejos moleculares. Pero, aunque hemos probado las partes del programa individualmente, todavía no se ha

validado el funcionamiento del STD-MaPa. Como deseamos conocer si la unión de todas las piezas del programa proporciona resultados fiables y cuantitativos, dedicaremos este apartado a la comprobación del STD-MaPa.

Para poner a prueba el programa, estudiaremos nuestros cinco complejos moleculares de referencia emulando un desarrollo estándar de análisis estructural basado en datos de STD-NMR. Como el programa está enfocado en la resolución analítica de las estructuras de los complejos moleculares, consideramos que, tanto la estructura del ligando como la disposición de los aminoácidos localizados en el sitio de unión han sido halladas a través de otras técnicas experimentales. Por tanto, los ligandos son separados de sus respectivas dianas para, posteriormente acoplarlos mediante el programa STD-MaPa. La exploración espacial se va a realizar en un paralelepípedo cúbico con 25 Å de aristas centrado en la cavidad de las proteínas. Como recordaremos, el STD-MaPa necesita generar al menos 5 modos de unión para lograr crear la estructura real. Por consiguiente, en una primera prueba, se generarán solo cinco, Y con esto, disminuirémos el tiempo del cálculo. Si el STD-NMR no es capaz de crear el 100% de estructuras reales, la cantidad de acoplamientos se irá incrementando paulatinamente. En el estudio, la optimización paramétrica se realizará para todos los rangos de valores permitidos por la última versión del programa. Finalmente, las intensidades de STD-NMR de cada unión serán evaluadas con los factores R_e y R_t . Con este procedimiento no se tiene en cuenta el resto de los factores R-NOE, porque el número de estructuras a evaluar no es mayor a cinco. Por tanto, R_{tm} , R_{twm} , R_{ms} y R_{msw} no son útiles.

Los cálculos se realizaron en una máquina Linux, con la distribución Ubuntu 16.04, un procesador Intel *Core 2 Duo* a 2.4 GHz y 4Gb de RAM. Los datos se introdujeron mediante la interfaz gráfica y los cálculos corrieron en paralelo mediante el uso de los dos núcleos del procesador. Los resultados fueron visualizados con el STD-MaPa y graficados con el QtiPlot.

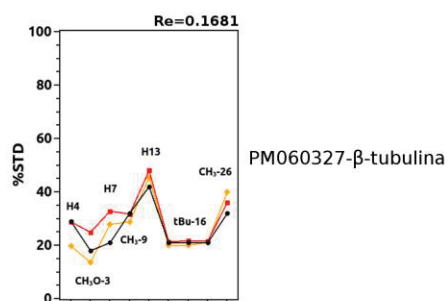
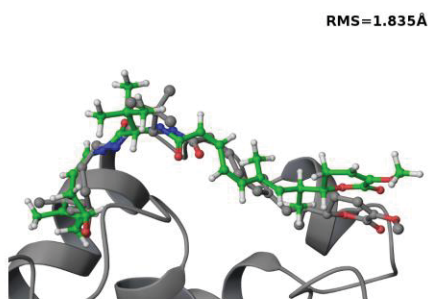
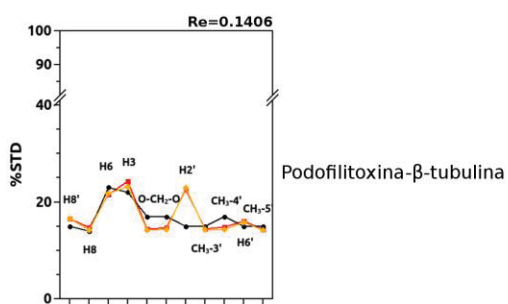
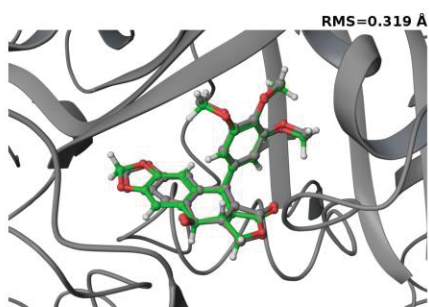
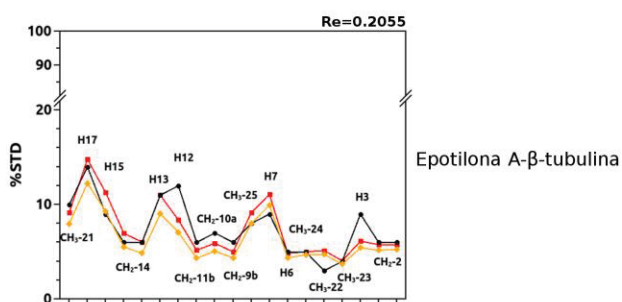
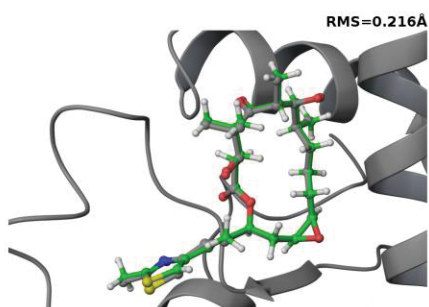
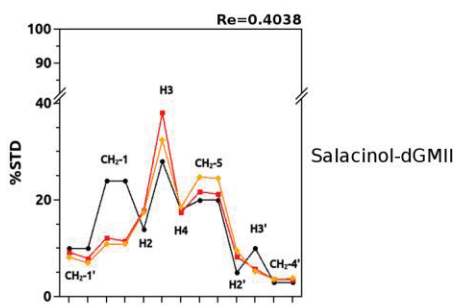
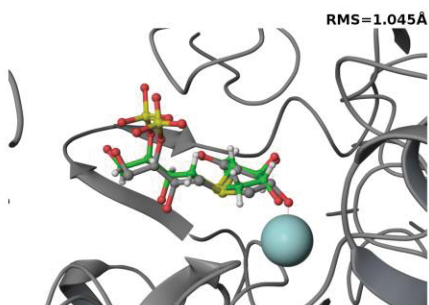
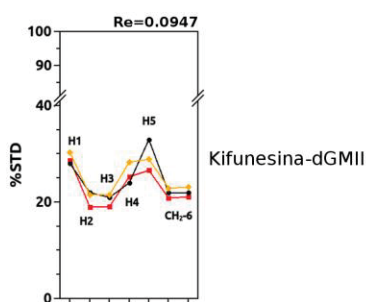
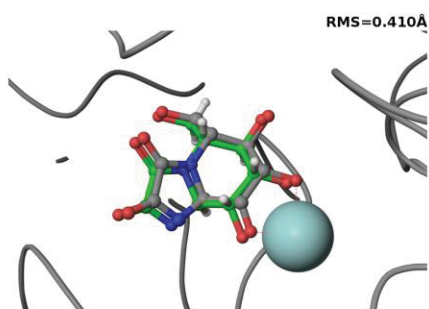


Figura 5.2.11. Representación de las coordenadas de la unión obtenidas con el STD-MaPa. A la izquierda, las estructuras de referencia (gris) y las obtenidas con el STD-MaPa (verde) se encuentran superpuestas. A la derecha, los datos de STD-NMR experimentales (negro), los datos de STD-NMR calculados a partir de las estructuras de referencia (rojo) y los datos arrojados por el STD-MaPa (naranja) son mostrados.

A partir del análisis de los resultados se extrae que los dos factores escogen el mismo modo de unión en los cinco complejos moleculares. Las estructuras de referencia y las obtenidas con el STD-MaPa para cada complejo molecular se han superpuesto en la Figura 5.2.11. Ahí, se puede observar con claridad como a partir del STD-MaPa se han resuelto correctamente las cinco estructuras. Si comparamos las intensidades experimentales, las calculadas a partir de las estructuras de referencia y las registradas por el programa, verificamos como las tres distribuciones siguen el mismo patrón.

En el cálculo del RMSD para todos los complejos moleculares se obtuvo que entre las coordenadas de referencia y las generadas por el STD-MaPa existen unas diferencias menores a 2 Å, siendo el PM060327 el que tiene mayor discrepancia con un RMSD de 1,835 Å. En este complejo, la estructura calculada por medio del STD-MaPa es lineal y exhibe los mismos puntos de anclaje encontrados en la estructura cristalográfica. Solo el heterociclo de lactona esboza una ligera modificación, cambio que produce un aumento de la exposición del metóxido al solvente.

Ésta diferencia en la orientación repercute en las intensidades de los protones H7, CH₃O-3 y H4, las cuales se redujeron entre un 5 y un 15%. El cambio conformacional tiene su mayor efecto sobre las intensidad de CH₃O-3, reduciendo su valor un ~15% por quedar más expuesto al solvente. Curiosamente, la reducción sufrida por los protones de CH₃O-3 ha corregido su valor, acercándolo más al experimental. En los datos de STD-NMR del PM060327 es fácil identificar los protones que se encuentran orientados hacia

el solvente por su menor intensidad. Este es el caso del tBu-16, el H7 y CH₃O-3. Estos resultados nos inducen a suponer que las interacciones con la Lys102 y la Asn105 no retienen totalmente la lactona, quedando más expuesta al solvente. Finalmente, se puede concluir que según los datos expuestos anteriormente, la estructura resuelta mediante el STD-MaPa describe mejor los datos de STD-NMR que las coordenadas cristalográficas.

En los últimos estudios hemos analizado los diferentes aspectos del programa que hemos desarrollado. A partir de los resultados, podemos garantizar que la metodología desarrollada, bajo ciertas condiciones, asegura la resolución de la estructura de complejos proteína-ligando a partir de datos de STD-NMR. En este algoritmo, se implementaron seis factores R-NOE para evaluar un conjunto de orientaciones intermoleculares en la unión; de ellos, cinco fueron diseñados para mejorar la robustez, sensibilidad y selectividad del factor R-NOE existente. Con estos factores se ha llegado a identificar cinco estructuras de complejos proteína-ligando escogidas como referencia. Finalmente, el programa STD-MaPa fue utilizado en la resolución estructural de estos complejos moleculares y, a partir de los resultados, hemos identificado una diferencia entre la estructura cristalográfica y los datos de STD-NMR para la lactona terminal del PM060327 en la unión a la β -tubulina. De esta forma, los datos de RMN llegan a describir una conformación menos anclada al surco de la proteína, que puede repercutir en la entropía de unión del sistema.

5.2.7 Cálculo de múltiples modos de unión con el STD-MaPa

La aproximación del CORCEMA-ST puede ser utilizada en la caracterización de múltiples modos de unión (Angulo, Díaz et al. 2008), siempre que se consideren las intensidades obtenidas del experimento de STD-NMR como la suma de las intensidades acumuladas a partir de cada modo de unión.

$$STD_{exp} = \frac{I_0 - I_{sat}^{exp}}{I_0} = \frac{\sum_{i=1}^m (I_0 - I_{sat}^i)}{I_0} = \sum_{i=1}^m STD_i \quad 5.2.1$$

Así que el cálculo de cada modo de unión puede realizarse por separado considerando constantes de asociación y concentraciones iniciales específicas para cada unión. Debido a la ambigüedad en las densidades electrónicas en las cavidades de β -tubulina, no se logró definir el número total de asociaciones del compuesto. Por este motivo se empezó por describir dos modos de asociación con diferentes proporciones. Si esta descripción no ajustase correctamente el patrón suma de intensidades de STD-NMR, se procedería a caracterizar un número creciente de modos de unión no covalentes.

El STD-MaPa puede ser utilizado en la resolución estructural de complejos con múltiples modos de unión, siempre que se conozcan las concentraciones del complejo molecular, la proteína y el ligando que están involucrados por cada modo de unión. La aproximación CORCEMA adaptada a los experimentos de STD-NMR no fue desarrollada para caracterizar múltiples asociaciones, por lo tanto fue necesario determinar que fracción de las concentraciones iniciales del ligando y la proteína están involucradas por cada unión.

En el 2008, el cálculo de las concentraciones iniciales por modo de unión fue realizado para un sistema molecular que exhibía dos modos de unión y consideraba una relación fija de concentraciones (Angulo, Díaz et al. 2008). En nuestro estudio, todavía no se conoce el número de uniones que están describiendo los datos de STD-NMR y la proporción en la que se encuentran. Por lo que, si se desea realizar una caracterización rigurosa mediante la aproximación de múltiples modos de unión, las proporciones relativas entre uniones y el total de uniones deberían ser optimizadas. Por la dificultad que conlleva la resolución de las relaciones de concentración para cada condición, en este trabajo se desarrollará una ecuación generalizada que permita obtener

el valor de las concentraciones iniciales por modo de unión, independientemente del número de uniones y de sus proporciones.

5.2.7.1 *Desarrollo de las ecuaciones generales de concentraciones para la descripción de múltiples modos de unión con el CORCEMA (EGC-CORCEMA)*

En la aproximación de múltiples modos de unión del CORCEMA-ST, se considera que los intercambios químicos entre el estado libre y cada modo de unión son independientes. Por lo tanto, si el ligando se asocia a la proteína con más de una orientación, se puede considerar que la relajación de la magnetización entre los diferentes modos de asociación es despreciable, y por tanto, la suma de las saturaciones correspondientes a cada modo de unión es igual a la saturación experimental. Mediante esta consideración, se obtuvieron las concentraciones iniciales y teóricas de cada intercambio químico (concentraciones iniciales ficticias) a partir de las concentraciones iniciales utilizadas en el experimento de STD-NMR.

Para llegar a una ecuación generalizada que permita obtener las concentraciones iniciales ficticias, se partió de la relación de concentraciones.

$$[L] + [P] \rightleftharpoons \sum_{i=1}^m [PL]_i \quad 5.2.2$$

En la relación de concentraciones descrita en 5.2.2, m es el total de orientaciones que adquiere el ligando al unirse a la proteína. En la aproximación de múltiples modos de unión, el equilibrio químico de cada asociación se describe individualmente y para ello los procesos tienen que ser independientes entre sí.

$$K_a^i = \frac{[PL_i]}{[P][L]} \quad 5.2.3$$

En la ecuación 5.2.3, K_a^i es la constante de asociación de la unión i-ésima y $[PL_i]$ es la concentración del complejo molecular cuando el ligando adopta la i-ésima orientación. Como la suma de las concentraciones parciales de las moléculas unidas y libres tiene que ser igual a la concentración inicial utilizada en el experimento, se puede relacionar la concentración inicial del ligando o la proteína con la combinación de sus respectivos estados de asociación en solución.

$$\begin{aligned} c_o(L) &= [L] + \sum_{i=1}^m [PL_i] \\ c_o(P) &= [P] + \sum_{i=1}^m [PL_i] \end{aligned} \quad 5.2.4$$

A partir de las últimas tres ecuaciones (ecuaciones 5.2.3 y 5.2.4) se pueden obtener las relaciones entre las constantes de asociaciones de cada modo de unión y las concentraciones de los complejos moleculares. Para esto, las relaciones entre las concentraciones fueron reordenadas, se despejaron las concentraciones de la proteína y el ligando libres, y se incluyeron los términos de proporcionalidad entre los pares de orientaciones $\left(y_i^k = \frac{[PL_i]}{[PL_k]}\right)$. Al realizar estos cambios, la constante de asociación se pudo reestructurar y dejar en función de la concentración del complejo molecular.

$$K_a^k = \frac{[PL_k]}{\left(c_0(P) - \left[\sum_{\substack{i=1 \\ i \neq k}}^m y_i^k + 1\right][PL_k]\right) \left(c_0(L) - \left[\sum_{\substack{i=1 \\ i \neq k}}^m y_i^k + 1\right][PL_k]\right)} \quad 5.2.5$$

En un primer paso se calculan las concentraciones de los complejos moleculares en el modo k-ésimo, para en un segundo paso calcular las concentraciones iniciales ficticias de cada estado de asociación. Por lo tanto, si redistribuimos la ecuación 5.2.5 se obtiene:

$$\begin{aligned} & \left(1 + \sum_{\substack{i=1 \\ i \neq k}}^m y_i^k\right)^2 K_a^k [PL_k]^2 \\ & - \left([c_0(P) + c_0(L)] \left(1 + \sum_{\substack{i=1 \\ i \neq k}}^m y_i^k\right) K_a^k + 1\right) [PL_k] \\ & + c_0(P)c_0(L)K_a^k = 0 \end{aligned} \quad 5.2.6$$

En la ecuación 5.2.6, los valores de $c_0(P)$, $c_0(L)$, y_i^k y K_a^k son constantes. Por lo tanto, solo habría que resolver los valores de $[PL_k]$. Los tres términos constantes de la ecuación 5.2.6 son sustituidos por tres variables constantes para simplificar la ecuación:

$$\begin{aligned}
 a_k &= \left(1 + \sum_{\substack{i=1 \\ i \neq k}}^m y_i^k\right)^2 K_a^k \\
 b_k &= -\left([c_0(P) + c_0(L)] \left(1 + \sum_{\substack{i=1 \\ i \neq k}}^m y_i^k\right) K_a^k + 1\right) \\
 c_k &= c_0(P)c_0(L)K_a^k
 \end{aligned}
 \tag{5.2.7}$$

Al sustituir los términos representados en 5.2.7 por tres constantes, quedaría una simple ecuación de segundo orden:

$$a_k[PL_k]^2 + b_k[PL_k] + c_k = 0 \tag{5.2.8}$$

Por lo tanto, la concentración del k-ésimo estado asociado se puede resolver como una ecuación de segundo orden.

$$[PL_k] = \frac{-b_k \pm \sqrt{b_k^2 - 4a_k c_k}}{2a_k} \tag{5.2.9}$$

Sin embargo, todas las soluciones en la resolución de la ecuación no tienen un significado físico. Por tanto, se necesitarían definir ciertas restricciones que permitiesen identificar la solución correcta. Por consiguiente, se plantearon tres

condiciones, y siempre que se cumplan las tres, se identificará la solución correcta.

Condiciones:

$$\forall [PL_i], [P], [L] > 0, \forall i \in \mathbb{N}, 1 \leq i \leq m$$

$$\text{cuando } c_0(L) > c_0(P); \text{ siempre } [L] > [P]$$

$$\text{si } c_0(L) \gg c_0(P); \text{ entonces } [P] \approx 0$$

Posterior a la determinación del aporte de cada modo de unión a la concentración total del ligando unido, se procedió a calcular las concentraciones iniciales ficticias que se deberían introducir en el programa. De este modo, se escribe nuevamente la ecuación de la constante de asociación, aunque en este caso solo se va a considerar un solo modo de unión.

$$K_a^k = \frac{[PL_k]}{(c_{0,k}'(P) - [PL_k])(c_{0,k}'(L) - [PL_k])} \quad 5.2.10$$

Las concentraciones iniciales marcadas con apóstrofes son las concentraciones iniciales ficticias que se adicionan al programa, para asegurar que la concentración del ligando unido ($[PL_k]$) se corresponda con los términos de proporcionalidad $\left(y_i^k = \frac{[PL_i]}{[PL_k]}\right)$ y se cumplan las relaciones descritas en las ecuaciones expuestas en 5.2.2 y 5.2.4. Así, se asegura que la fracción de ligando enlazado sea la misma en la solución real y en la separación ficticia de los diferentes uniones. Por otra parte, la relación entre las concentraciones experimentales de la proteína y el ligando no puede ser modificada, ya que

añadiría un error al cálculo de concentraciones. Para que esto no ocurra hay que asegurarse que la relación $\frac{c_{0,k}'(L)}{c_{0,k}'(P)}$ siempre sea igual a $\frac{c_0(L)}{c_0(P)}$. Por lo tanto, la concentración inicial ficticia de la proteína se sustituye por la relación de concentraciones experimentales ($c_{0,k}'(P) = \frac{c_0(P)}{c_0(L)} c_{0,k}'(L)$).

$$K_a^k = \frac{[PL_k]}{\left(\frac{c_0(P)}{c_0(L)} c_{0,k}'(L) - [PL_k]\right) (c_{0,k}'(L) - [PL_k])} \quad 5.2.11$$

Al igual que en la ecuación 5.2.5, también aquí se realiza una reestructuración para calcular $c_{0,k}'(L)$:

$$\begin{aligned} \frac{c_0(P)}{c_0(L)} K_a^k c_{0,k}'(L)^2 - \left(1 + \frac{c_0(P)}{c_0(L)}\right) K_a^k [PL_k] c_{0,k}'(L) \\ + (K_a^k [PL_k]^2 - [PL_k]) = 0 \end{aligned} \quad 5.2.12$$

y para simplificar la formulación, el conjunto de parámetros se agrupan en tres constantes,

$$\begin{aligned}
A_k &= \frac{c_0(P)}{c_0(L)} K_a^k \\
B_k &= -\left(1 + \frac{c_0(P)}{c_0(L)}\right) K_a^k [PL_k] \\
C_k &= K_a^k [PL_k]^2 - [PL_k]
\end{aligned}
\tag{5.2.13}$$

dando como resultado una simple ecuación de segundo orden.

$$A_k c_{0,k}'(L)^2 + B_k c_{0,k}'(L) + C_k = 0 \tag{5.2.14}$$

Por lo tanto, al igual que la concentración del complejo molecular en el k-ésimo estado de asociación, se calculó la concentración inicial ficticia del ligando para el estado k mediante la resolución de una ecuación de segundo orden. La solución tiene que respetar las condiciones expuestas anteriormente y las postuladas a continuación.

Condiciones:

$$\forall c_{0,k}'(P), c_{0,k}'(L), [P_k]', [L_k]' > 0, \forall k \in \mathbb{N}, 1 \leq k \leq m$$

$$c_{0,k}'(L) \geq [PL_k], \forall k \in \mathbb{N}, 1 \leq k \leq m$$

$$c_{0,k}'(P) \geq [PL_k], \forall k \in \mathbb{N}, 1 \leq k \leq m$$

$$\text{cuando } c_{0,k}'(L) > c_{0,k}'(P); \text{ siempre } [L_k]' > [P_k]' \text{ y } c_{0,k}'(L) > [PL_k]$$

Al resolver la concentración inicial y ficticia del ligando para cada estado k , la concentración inicial y ficticia de la proteína se determina mediante la relación de las concentraciones experimentales en el mismo estado k ($c_{0,k}'(P) = \frac{c_0(P)}{c_0(L)} c_{0,k}'(L)$).

A partir del desarrollo expuesto en el actual apartado se puede realizar el estudio de múltiples modos de unión. Al contar con ecuaciones generales, la automatización del proceso de resolución estructural de complejos moleculares con múltiples modos de unión a partir de datos de STD-NMR es posible.

Además, las fracciones de los ligandos ($f_{LB}^k = \frac{[PL_k]}{c_0(L)}$) y proteínas enlazadas ($f_{PB}^k = \frac{[PL_k]}{c_0(P)}$) pueden determinarse para cada modo de unión. El valor de f_{LB} es importante porque permite estimar el ancho de las bandas observadas de las señales del ligando y f_{PB} es usada para conocer la disponibilidad de los sitios libres en ensayos de competición (Meyer and Peters 2003).

5.3 Método híbrido de RMN y DRX para la resolución estructural de complejos proteína-ligando

5.3.1 Resolución de la estructura del complejo dactilolida-tubulina

El STD-MaPa se ha utilizado con éxito en la resolución de las estructuras de cinco complejos moleculares. Los resultados han sentado las bases para aplicar el programa en el estudio de complejos moleculares en sistemas reales. Así que nos planteamos resolver la estructura de la dactilolida sintética unida al heterodímero de α,β -tubulina.

La dactilolida es un probado agente estabilizante de microtúbulos que se une covalentemente a la β -tubulina mediante una cinética de reacción lenta (Field, Pera et al. 2012). Este compuesto ha exhibido en diversos estudios cierta toxicidad y la capacidad de evadir alguno de los mecanismos más comunes de resistencia tumoral (Cutignano, Bruno et al. 2001, Field, Pera et al. 2012). En el mismo estudio, mediante ensayos de espectrometría de masas acoplados a HPLC, se llegó a determinar que la dactilolida se une covalentemente tanto al residuo de HIS229 como a la ASN228. Ésta información se corroboró a partir de la estructura cristalográfica de la zampanolida unida al complejo T_2R . La zampanolida es un compuesto de la misma familia que la dactilolida que en el cristal se unió covalentemente al residuo de HIS229 de la β -tubulina (Prota, Bargsten et al. 2013).

Aunque la estructura del complejo ZMP- T_2R ha sido ya desvelada, debido a las características de la unión de la dactilolida la estructura del complejo molecular no se ha resuelto experimentalmente. Al unirse covalentemente la DAC y ZMP a la β -tubulina, el análisis estructural del proceso asociativo ayudaría en la búsqueda y diseño de nuevos fármacos que fuesen capaces de burlar la resistencia tumoral a fármacos. Por tanto, en este apartado vamos a hablar de una metodología híbrida, la cual diseñamos para resolver la estructura de

complejos moleculares combinando técnicas de RMN y DRX. Así, con la creación del método híbrido nos propusimos resolver la estructura del complejo molecular DAC-T₂R antes de reaccionar covalentemente. Ya que, con el análisis de las estructuras de ZMP-T₂R y DAC-T₂R, vamos a poder caracterizar el mecanismo molecular de la reacción covalente en el sitio del Taxol®.

El método híbrido fue diseñado porque, ni la DRX, ni las técnicas de RMN utilizadas aquí (TR-NOESY y STD-NMR), resuelven totalmente el sistema molecular de la dactilolida unida al heterodímero de α,β -tubulina. Con los datos de TR-NOESY y el STD-NMR solo podemos caracterizar la unión de la DAC en el sitio de β -tubulina mediante el STD-MaPa y algoritmos de exploración conformacional; mientras que, con la DRX y siguiendo el protocolo descrito en (Prota, Bargsten et al. 2013) solo se llega a resolver con alta resolución la estructura del complejo T₂R. Así que el método híbrido combina datos provenientes de los métodos cristalográficos o de crio-ME, TR-NOESY y STD-NMR (Figura 5.3.1) para obtener la información suficiente y así resolver todo el sistema de coordenadas del complejo molecular.

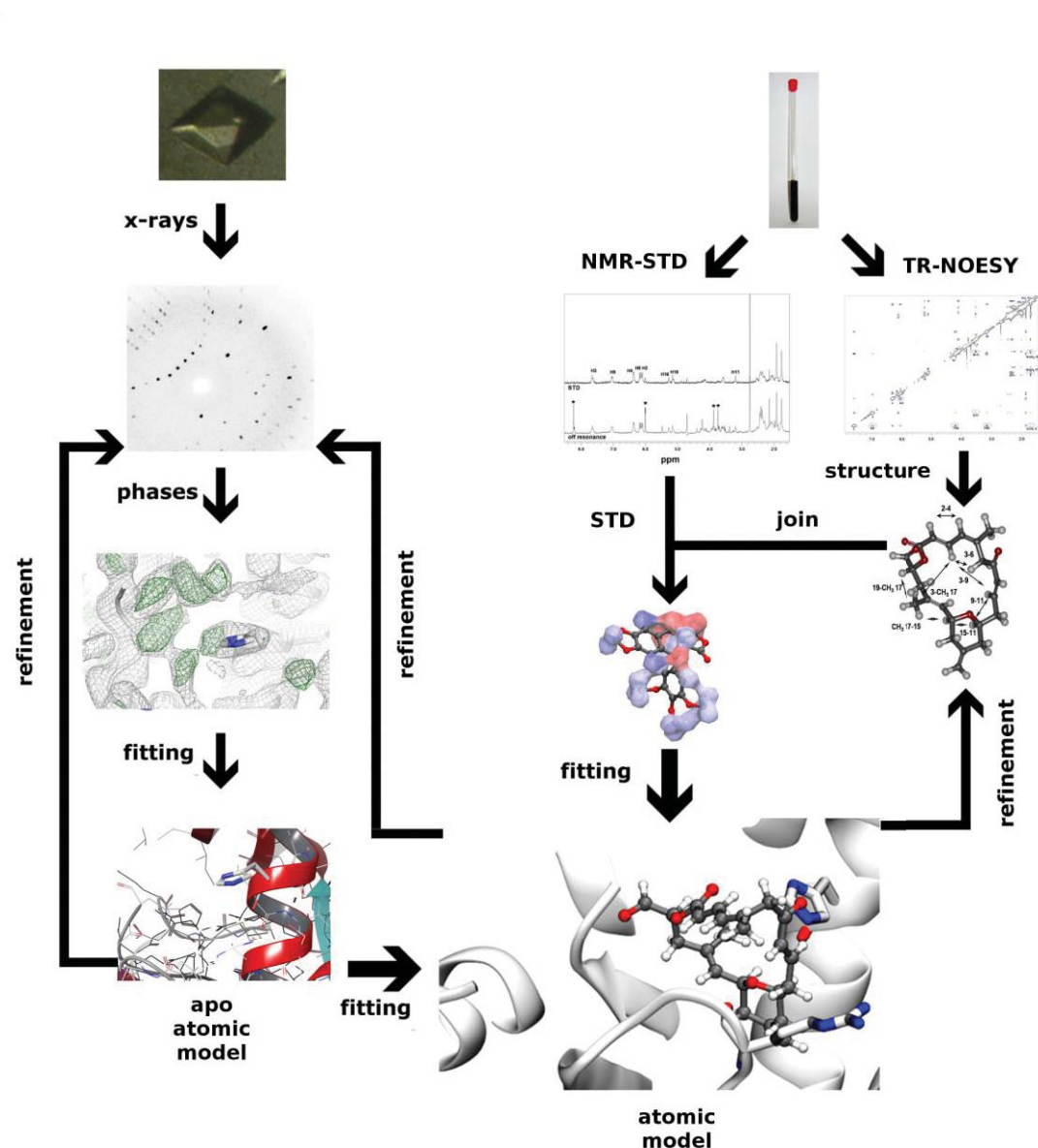


Figura 5.3.1. Esquema del método híbrido de resolución estructural de complejos moleculares basado en datos cristalográficos, TR-NOESY y STD-NMR.

Aunque en el método híbrido se utilizó la DRX para obtener las coordenadas de la proteína, la estructura de la proteína puede resolverse con cualquier técnica que describa parcialmente la unión de las moléculas de ligando y la proteína.

Entre las técnicas que se podrían usar se encuentran la crio-ME de alta resolución o técnicas de RMN dedicadas a la resolución de proteínas. Sin embargo, la resolución estructural de proteínas mediante técnicas de RMN es muy laboriosa y necesita personal muy especializado. Por otra parte, la DRX y crio-ME reducen su poder resolutivo en las regiones más flexibles. Así que las moléculas pequeñas que forman uniones no covalentes con las proteínas se resuelven con dificultad, porque son flexibles o tienen intercambios químicos rápidos. Por lo tanto, el método que se escoja dependerá de las condiciones y características del sistema que se va a estudiar.

Con este trabajo intentaremos proponer una alternativa que combine las virtudes de los diferentes métodos experimentales relatados. La composición más acertada que encontremos, partirá de mantener intactos los procedimientos experimentales estándares. A causa de que los laboratorios necesitan equipamiento y personal muy especializado, la metodología híbrida sería inviable si se realizan cambios radicales en los procedimientos experimentales estándares. Así que la solución idónea combinará solo la información estructural resuelta a partir de cada una de las técnicas, para complementar la ausencia total o parcial de información en el resto de técnicas.

Igualmente, el método híbrido conlleva ciertas dificultades experimentales. En el caso que nos ocupa, el complejo formado por la dactilolida y la α,β -tubulina se debe cristalizar el complejo molecular como un monocristal para que difracte adecuadamente y, además, los compuestos deberán disolverse adecuadamente en una solución deuterada que permita la medición en un espectrómetro de RMN. Pero cuando los métodos cristalográficos y de RMN no resuelven la estructura individualmente, y el sistema en estudio cumple con las condiciones experimentales necesarias para estudiarlos parcialmente con las dos aproximaciones, su resolución se asegura mediante el método híbrido.

El método híbrido está formado por tres técnicas que se encargarán de resolver y refinar la estructura de la proteína (DRX), el ligando (TR-NOESY) y el acoplamiento entre ellos (STD-NMR). Los resultados de cada técnica tienen

que poder usarse en el resto, así que las preparaciones de las diferentes muestras no podrán modificar la estructura del complejo molecular. Para utilizar el método híbrido es necesario el programa STD-MaPa, porque la metodología implementada en él se encargará de combinar las estructuras obtenidas por DRX y TR-NOESY. Anteriormente al diseño del STD-MaPa, las tres metodologías no se habrían podido combinar satisfactoriamente, porque las metodologías de acoplamiento molecular basadas en funciones de energía no aseguraban la descripción correcta de los datos de STD-NMR experimentales.

El STD-MaPa se utiliza para optimizar las variables químico-físicas y al mismo tiempo escoger las orientaciones a partir de las primeras posiciones de la lista de R_e , R_t y R_{tm} (ecuaciones 1.2.10, 5.1.1 y 5.1.2). En este paso, se escogen varias posiciones para tener más información en el paso de refinamiento mediante las técnicas de cristalografía y TR-NOESY (Figura 5.3.1). Posterior al paso de acoplamiento y antes del refinamiento con las otras técnicas, las estructuras se someterán a un refinamiento conformacional a través del STD-MaPa o de SICO (Jayalakshmi, Biet et al. 2004).

Las conformaciones del ligando obtenidas del STD-MaPa o SICO son refinadas mediante una minimización conformacional local basándonos en las restricciones de TR-NOESY y reinsertadas en el STD-MaPa. A partir de las estructuras del STD-MaPa se recalculan las densidades electrónicas del cristal (técnica de DRX). Al estar el ligando localizado en el sitio y basándonos en la suposición de que las dos técnicas están describiendo el mismo acoplamiento, se espera con las nuevas estructuras que aparezcan nuevas densidades electrónicas y además que puedan explicar los mapas de densidades de contorno positivo (regiones de la estructura real que no están presentes en el modelo) encontrados en la apo-estructura del cristal. Finalmente, las estructuras cristalográficas refinadas son reinsertadas en el STD-MaPa, las constantes químico-físicas son nuevamente optimizadas y los factores R_t y R_e calculados. El procedimiento es autoconsistente y culminará cuando se

encuentre la estructura molecular que mejor ajuste los datos de las tres técnicas.

5.3.2 Cristalografía del complejo dactilolida-tubulina

Por lo tanto, una primera versión de la estructura de la apo-proteína con una resolución de 2,5 Å fue resuelta siguiendo el protocolo descrito para la unión de la zampanolida y la epotilona-A con el complejo tetramérico T₂R (Prota, Bargsten et al. 2013) (Figura 5.3.2-A). Los datos cristalográficos son mostrados en la Tabla 5.3.1. A partir del cristal, no se pudo resolver la dactilolida en las cavidades de β1- y β2-tubulina. Sin embargo, al analizar los mapas de densidades en las cavidades de las proteínas (Figura 5.3.2-B), encontramos densidades que no pertenecen ni a la proteína, ni a moléculas de agua. Así que las intensidades de difracción experimentales (F_{obs} , φ_{calc}) están describiendo una estructura en la cavidad de β-tubulina que todavía no hemos incluido en el modelo (ver los mapas de diferencias de densidad de contorno positivo ilustrados como un enrejado verde en la Figura 5.3.2-B). Curiosamente, el mapa de diferencias de densidad ($F_{obs} - F_{calc}$, φ_{calc}) exhibe una mayor densidad electrónica en la cavidad de β1-tubulina que en β2-tubulina.

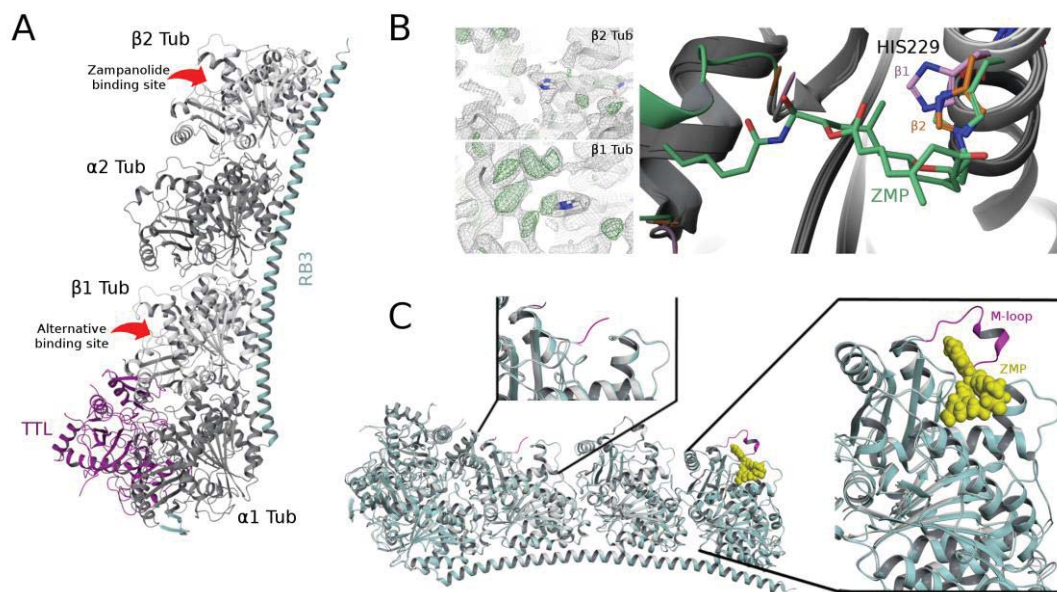


Figura 5.3.2. (A) Representación de la estructura cristalográfica del complejo dactilolida- T_2R . La dactilolida no se representa porque no existe densidad electrónica suficiente para poder describir su estructura. (B) A la izquierda, representación de la densidad electrónica en $\beta 1$ - y $\beta 2$ -tubulina. En verde, el mapa de diferencia de contorno positivo de densidades electrónicas ($F_{obs} - F_{calc}$, φ_{calc}). En gris, la densidad electrónica observada (F_{obs} , φ_{calc}). A la derecha, superposición de la estructura de $\beta 1$ - y $\beta 2$ -tubulina del complejo de dactilolida- T_2R y la $\beta 2$ -tubulina del complejo de zampanolida- T_2R (PDB: 4I4T). En verde se representa el lazo-M, la zampanolida y la HIS229. (C) Superposición del complejo T_2R cristalizado con la dactilolida (en gris) y la zampanolida (azul claro) respectivamente. En la ilustración se procedió a aumentar los sitios de unión localizados en $\beta 1$ - y $\beta 2$ -tubulina. La ZMP está representada como esferas de Van der Waals amarillas y el lazo-M está resaltado con el color morado.

Tabla 5.3.1. Tabla de datos cristalográficos del complejo dactilolida-T₂R.

	Dactylolide
Data collection	
Space group	P212121
Cell dimensions	
a, b, c (Å)	104.6, 157.7, 180.2
α, β, γ (°)	90, 90, 90
Resolution (Å)*	47.8-2.1 (2.15-2.1)
No. reflections	173479 (12676)
R_{meas} (%)	13.8 (266.0)
R_{pim}	4.2 (76.4)
I / σI	16.1 (1.1)
Completeness (%)	99.9 (100.0)
Redundancy	13.7 (13.6)
Refinement	
Resolution (Å)	47.8-2.1
No. unique reflections	173479
R_{work} / R_{free} (%)	20.0 / 22.4
Average B-factors (Å²)	
Complex	50.6
Solvent	49.5
Wilson B-factor	40.9
Root mean square deviation from ideality	
Bond length (Å)	0.005
Bond angles (°)	0.803
Ramachandran statistics^c	
Favored regions (%)	97.35
Allowed regions (%)	2.50
Outliers (%)	0.15

*Highest-resolution shell is shown in parentheses. ^cAs defined by MolProbity(Davis, Murray et al. 2004)

Antes de acoplar los resultados cristalográficos con los de RMN se debe comprobar que la estructura del cristal describe la unión con el compuesto. Para ello, se verifica la existencia de evidencia estructural que indique que la cavidad está ocupada por el compuesto. En nuestro caso, comprobamos que hubiera densidad electrónica en la cavidad que no perteneciese a la proteína, y que los aminoácidos que forman parte de la cavidad hubiesen cambiado su orientación respecto a la estructura de T₂R cristalizada en ausencia de cualquier tipo de ligando (PDB: 4IIJ) (Prota, Magiera et al. 2013) y se asemeje a la estructura del complejo ZMP-T₂R (PDB: 4I4T) (Prota, Bargsten et al. 2013).

Cuando se compararon las orientaciones de los aminoácidos en los dos sitios de la apo-estructura del cristal con dactilolida (β 1- y β 2- tubulina), con la apo-estructura obtenida en ausencia de ligando y en presencia de ZMP, se comprobó que tanto en la β 2- como en la β 1- tubulina del cristal con DAC, la HIS229 se encontrara orientada hacia el interior del sitio preparada para el enlace al igual que en el cristal con la ZMP (Figura 5.3.2-B). Si la estructura cristalográfica no estuviera describiendo la unión del compuesto, la HIS229 se orientaría hacia el exterior del sitio, como ocurre con la estructura en ausencia de ligando.

Si la cavidad de la tubulina se encuentra desocupada, la HIS229 se orienta en dirección opuesta al sitio; mientras que cuando el sitio luminal está ocupado, la histidina se orienta hacia el interior del sitio de unión para formar parte de la interacción con el ligando. En el caso de la zampanolida y la dactilolida, la HIS229 se enlaza covalentemente con el carbonilo α,β -insaturado del macrociclo de ambas moléculas mediante una adición de Michael. Por tanto, la orientación de la histidina 229 nos brinda una clara evidencia de la presencia de dactilolida en ambos sitios de unión (Figura 5.3.2-B). En el Panel B de la Figura 5.3.2 se superpusieron las cavidades de β 1- y β 2-tubulina del complejo dactilolida-T₂R junto con el sitio de unión de la zampanolida-T₂R. En ésta, se visualiza claramente la orientación adquirida por la HIS229 en los tres casos, estando el residuo favorablemente orientado hacia la formación del enlace

covalente. En el caso particular de β 1-tubulina, la cadena lateral adquiere una ligera desviación torsional que sitúa el residuo hacia el centro del macrociclo de lactona.

Al caracterizar la unión de la dactilolida y zampanolida por métodos bioquímicos se comentó que la cadena lateral N-acil hemiaminal de la zampanolida estabilizaba la estructura del lazo-M (Figura 5.3.2-C). Este efecto también ocurre con la cadena lateral de la epotilona-A y sus derivados (Prota, Bargsten et al. 2013, Jantsch, Nieto et al. 2016). Sin embargo, cuando los compuestos no tiene cadena lateral o en ausencia de compuestos en el sitio interno, el lazo-M se desestructura y no se puede determinar por técnicas cristalográficas. Así que, si se desea obtener la estructura 3D del lazo mediante técnicas de DRX, el sitio debe estar ocupado por ligandos que presenten interacciones directas con el lazo-M. Pero, a partir de nuestro cristal, no obtenemos densidad electrónica en la región del lazo-M (Figura 5.3.2-C). Estos resultados concuerdan con la ausencia de cadena lateral en la estructura de la dactilolida y los ensayos cinéticos previos donde se estudiaba la asociación de la DAC.

5.3.3 Medición de TR-NOESY y STD-NMR del complejo dactilolida-tubulina

La ausencia de densidad electrónica y la baja resolución en las cavidades de las dos β -tubulinas impidieron la localización del ligando, así que solo se obtuvieron las posiciones atómicas de la apo-estructura del complejo T₂R en presencia de la DAC. Por consiguiente, la conformación bioactiva y la orientación de la DAC en el heterodímero de α,β -tubulina tiene que determinarse mediante TR-NOESY y STD-NMR. Por lo tanto, se preparó una solución deuterada con 300 μ M de DAC y 10 μ M de tubulina bajo condiciones no polimerizantes en presencia de DMSO-d₆ (ver apartado 4.5), y se realizaron

los experimentos de STD-NMR y TR-NOESY como se describió en el apartado 4.6.

Los espectros se adquirieron en un espectrómetro Bruker AVANCE de 600 MHz a 298 K (Figura 5.3.3-A y C). Los datos extraídos fueron tratados con el programa TopSpin de Bruker para obtener los espectros de TR-NOESY y STD-NMR. En cada caso, se identificaron los picos y se extrajeron las intensidades (Figura 5.3.3-B y D). Aunque la resolución estructural aumenta con el número de datos extraídos por espectro, para describir la estructura cuantitativamente, solo se tuvieron en cuenta los picos no solapados y que tuvieran una SNR > 10. Así, aseguramos que solo los picos con mayor resolución espectral y menos incertidumbre sean utilizados en el estudio estructural.

Aunque se tengan estos cuidados, para que la extracción de los datos sea analítica hay que monitorizar las dos asociaciones (la unión reversible y la irreversible). Por lo que, en el transcurso de los experimentos, se fueron adquirieron espectros de RMN-1D H^1 para controlar la asociación irreversible. La reacción covalente de la dactilolida es lenta ($k = 0,12 \text{ hr}^{-1}$), así que las técnicas de TR-NOESY y STD-NMR solo describen la reacción reversible ($k = 1 \text{ hr}^{-1}$). Sin embargo, al comparar las intensidades de los espectros de RMN-1D H^1 , pudimos calcular la velocidad de disminución promedio de la señales y, con ello, la velocidad de la unión irreversible. Finalmente, la velocidad de disminución de las intensidades obtenida por RMN fue similar a la velocidad de la reacción irreversible obtenida por los ensayos de fluorescencia.



Figura 5.3.3. Caracterización mediante STD-NMR y TR-NOESY de la interacción dactilolida- α,β -tubulina. (A) Espectro unidimensional protónico y espectro de STD-NMR de la dactilolida en presencia de la α,β -tubulina. (B) Representación de la estructura 3D de la dactilolida y de las intensidades relativas de STD-NMR. Las esferas ubicadas en las coordenadas protónicas muestran las intensidades relativas de STD-NMR extraídas del espectro. Las esferas azules oscuro representan los protones sin datos de STD-NMR. (C) Espectro bidimensional de TR-NOESY. Las señales de los Tr-NOE son resaltadas. (D) Superposición de seis conformaciones de dactilolida que cumplen con las restricciones impuestas por las señales del espectro de TR-NOESY. Las fuerzas de restricción utilizadas en la exploración conformacional están representadas con barras amarillas (50 kJ/molÅ²), naranjas (100 kJ/molÅ²) y rojas (200 kJ/molÅ²).

Si se desea resolver analíticamente las coordenadas atómicas del complejo DAC-T₂R, la información aportada por la DRX y RMN tendrá que igualar los grados de libertad del sistema molecular en un espacio euclidiano. Como las coordenadas de la proteína son obtenidas por DRX, las técnicas de RMN tendrán al menos que igualar 28 grados de libertad debidos a 22 grados torsionales, 3 traslacionales y 3 rotacionales. Si calculamos el total de intensidades que se extrajeron de los experimentos de TR-NOESY y de STD-NMR, nos daremos cuenta que contamos con 28 datos experimentales para describir la conformación y la orientación del ligando en el sitio de unión. Así que contamos con suficiente información para describir analíticamente la estructura del complejo molecular.

En el siguiente paso del protocolo híbrido, la estructura 3D de la dactilolida se resolverá transformando los datos de TR-NOESY en restricciones espaciales. El proceso consta de varios pasos que guían a los algoritmos de modelado molecular hacia la obtención de la estructura bioactiva y, en todos los pasos, se utilizará el algoritmo sistemático de pseudoMonte-Carlos implementado en el Macromodel (Mohamadi, Richards et al. 1990) para realizar la exploración conformacional. El método se deriva del método de Monte-Carlo de coordenadas internas (Chang, Guida et al. 1989), y se diferencia de éste en el proceso de exploración; ya que, en el método de pseudoMonte Carlo, la búsqueda aleatoria se sustituye por una búsqueda sistemática para mejorar la eficiencia del algoritmo.

Una primera exploración fue realizada sin restricciones atómicas para escoger por simple inspección visual las conformaciones que cumpliesen con las restricciones impuestas por los Tr-NOEs. En una segunda exploración conformacional, un potencial armónico es impuesto para restringir la distancia de los pares de átomos pesados enlazados a protones con Tr-NOEs. En este paso, 14 Tr-NOEs son utilizados para crear tres niveles de restricciones según sus intensidades; niveles que se dividen según la evidencia experimental. Las intensidades se dividieron, según su valor relativo, en intensidades débiles,

medias y fuertes; asignando una fuerza restrictiva de 50, 100 y 200 kJ/molÅ² respectivamente. Adicionalmente, se asigna un potencial de fondo para contrarrestar las fuerzas de restricción cuando las distancias interatómicas son menores a 6 Å. Los grupos químicos que no están relacionados por los Tr-NOEs se dejan evolucionar libremente bajo el campo de fuerzas OPLS2005.

La Figura 5.3.3-D muestra en código de colores los potenciales armónicos comentados anteriormente. En la imagen se puede observar cómo las restricciones están homogéneamente distribuidas por el macrociclo y las zonas flexibles de la molécula. Por este motivo, creemos que hay suficientes Tr-NOEs para predecir correctamente toda la estructura bioactiva de la dactilolida.

Seis confórmeros fueron aceptados por cumplir las restricciones impuestas en el modelado molecular guiado por los Tr-NOEs. Se superpusieron las estructuras de los seis confórmeros sobre la estructura de la zampanolida para determinar las diferencias estructurales en el macrociclo debidas a la técnica experimental. A los 7 macrociclos se les calculó el RMSD y se constató que la técnica de TR-NOESY describía la misma conformación bioactiva que fue descrita para la zampanolida por cristalografía.

5.3.4 Resolución de la estructura del complejo con STD-NMR

Una vez resuelta la estructura de la proteína y de la dactilolida, se puede proceder a calcular la orientación que tiene el ligando en la cavidad de β -tubulina. La unión de la DAC al heterodímero de α,β -tubulina fue caracterizada mediante la técnica de STD-NMR (Figura 5.3.3-A). Al analizar el espectro de STD-NMR se observó que la saturación de los protones fue homogénea sobre todo el esqueleto molecular, lo que sugiere que toda la molécula está involucrada en la unión. Al extraer los datos de STD-NMR siguiendo las restricciones comentadas en el capítulo 4, solo los protones pertenecientes al anillo 4-metilen-tetrahidropirano quedaron descartados en nuestro estudio

(Figura 5.3.3-B). Sin embargo, la ausencia de información en esta zona no impide la determinación de la unión, ya que la estructura de este anillo es rígida y está ligada químicamente al resto de la molécula. Por lo tanto, tendríamos información estructural suficiente y con calidad para proceder con la resolución total de la estructura del complejo de dactilolida- α,β -tubulina. En este paso, el programa STD-MaPa es utilizado para resolver la unión mediante la combinación de la estructura de la apo-tubulina, los seis conformeros y los datos de STD-NMR.

El acoplamiento de la dactilolida se obtuvo con el programa STD-MaPa, y para realizar este trabajo se disponía de toda la información necesaria; ya que se conocía la estructura de la proteína, los conformeros del ligando y los datos de STD-NMR. Además, la unión covalente de la zampanolida estaba caracterizada por DRX y la estructura de este compuesto es muy parecida a la dactilolida.

La estructura cristalográfica del complejo de dactilolida-tubulina fue resuelta con el complejo T₂R; mientras que los datos de RMN fueron obtenidos a partir de una solución no polimerizante del heterodímero de α,β -tubulina. Sin embargo, las diferencias estructurales entre el complejo T₂R y un heterodímero en solución no afectan la conformación de los residuos pertenecientes al sitio activo. La disposición geométrica del sitio de unión en β -tubulina respecto a la localización de la proteína RB3 y la enzima tubulina-tirosina ligasa (TTL) no inducen interacciones directas o alostéricas que produzcan cambios en la estructura del sitio de unión. Por otra parte, el acoplamiento interdimérico tampoco afecta al sitio del Taxol®. Si fuera así, se hubiesen encontrado evidencias claras en toda la cavidad al superponer β 1- y β 2-tubulina (Figura 5.3.2-B).

Por último, el lazo-M no se encuentra estabilizado en la estructura cristalográfica y esto le proporciona cierta flexibilidad (Figura 5.3.2-A,C). La flexibilidad del lazo-M se debe a la ausencia de una cadena lateral en la dactilolida que interaccione con esta región de la proteína. Así que, en nuestro estudio asumimos que las interacciones dipolares entre la dactilolida y el lazo-

M son depreciables, al igual que la transferencia de saturación efectiva desde esa región de la proteína sobre el ligando. Por lo tanto, el estudio se realizó en ausencia del lazo-M.

Las condiciones escogidas para el cálculo con el STD-MaPa son las mismas que las usadas en el apartado 5.2.6. En este caso, se crearon dos cajas cúbicas centradas en la cavidad de $\beta 1$ - y $\beta 2$ -tubulina, se utilizaron los seis conformeros y por cada uno, 5 orientaciones fueron generadas. El tiempo de correlación de la proteína fue fijado en 40 ns. Este valor se obtuvo en la optimización paramétrica realizada en los complejos de epotilona-A y podofilotoxina (Tabla 6.3.1). El resto de constantes químico-físicas fueron optimizadas a partir de la exploración de todo el rango de constantes permitido por el programa. Para este trabajo, los factores R_e y R_t fueron utilizados con el fin de evaluar las estructuras resultantes. En total se realizaron doce cálculos partiendo de los seis conformeros y las dos β -tubulinas.

El STD-MaPa arrojó una lista de estructuras que en los primeros puestos combinaba conformeros, constantes de equilibrios y velocidades de disociación. El tiempo de correlación del compuesto libre fue el único parámetro que se determinó por consenso (Tabla 5.3.2); mientras que las velocidades de disociación en $\beta 1$ - y $\beta 2$ -tubulina se mantuvieron en un rango de 10 s^{-1} a 100 s^{-1} .

Tabla 5.3.2. Resultados del STD-MaPa de las dos primeras posiciones de las listas creadas por el factor R_e en $\beta 1$ - y $\beta 2$ - tubulina.

Isotipo	Factor R_e	$K_a \text{ (M}^{-1}\text{)}$	$k_{\text{off}} \text{ (s}^{-1}\text{)}$	$\tau_c^l \text{ (ns)}$
$\beta 1$	0,29	10^4	10	0,75
$\beta 1$	0,38	10^8	100	0,75
$\beta 2$	0,35	10^5	10	0,75
$\beta 2$	0,46	10^4	10	0,75

Al comparar la estructura de las dos primeras posiciones de la lista de R_e en $\beta 1$ - y $\beta 2$ -tubulina, se encontró una elevada similitud entre las disposiciones espaciales de las uniones en ambas cavidades (Figura 5.3.4). Los cálculos de los RMSD entre las uniones más favorables ($R_e(\beta 1)=0,29$ y $R_e(\beta 2)=0,35$) y entre las siguientes ($R_e(\beta 1)=0,38$ y $R_e(\beta 2)=0,46$) fueron menores a 2 Å, probando que el programa STD-MaPa escogió las mismas orientaciones de DAC para las dos β -tubulinas. Pero, dada la pequeña diferencia estructural en los residuos de ambos sitios, las conformaciones de DAC entre $\beta 1$ - y $\beta 2$ -tubulina difirieron ligeramente.

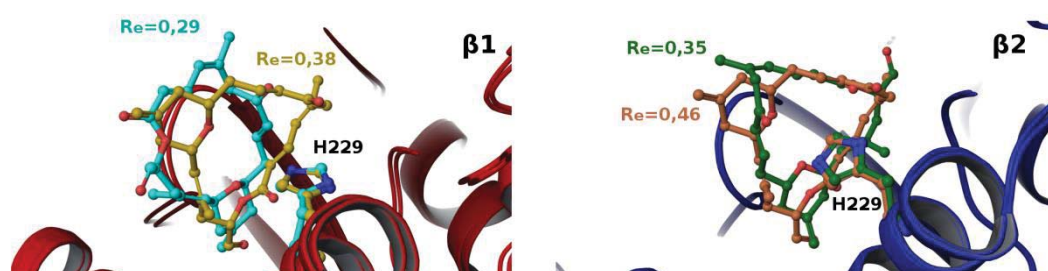


Figura 5.3.4. Superposición de las poses de unión determinadas por el STD-MaPa para el complejo de dactilolida y β -tubulina. Las dos primeras orientaciones de la lista del factor R_e son representadas. En la izquierda, representación del acoplamiento 1 ($R_e(\beta 1)=0,29$) (azul claro) y 2 ($R_e(\beta 2)=0,35$) (amarillo) en la $\beta 1$ -tubulina del complejo T_2R . En la derecha, representación del acoplamiento 1 ($R_e(\beta 1)=0,38$) (verde) y 2 ($R_e(\beta 2)=0,46$) (marrón) en la $\beta 2$ -tubulina del complejo T_2R .

Al comparar los patrones de intensidades nos percatamos que los modos de unión no describían correctamente el patrón de intensidades de STD-NMR (Figura 5.3.5-A). Ninguna orientación llega a describir completamente el mapa de saturación obtenido experimentalmente. Por tanto, los valores de R_e y R_t tienden a ser elevados. Según el factor R_e (Tabla 5.3.2), la orientación del

ligando y los residuos de β 1-tubulina se adaptan mejor a los datos de STD-NMR que la orientación en β 2-tubulina. En cualquier caso, las uniones describieron correctamente solo la mitad del patrón de intensidades experimentales, ya que, en el cálculo de las intensidades de STD-NMR, la otra mitad de los protones se redujo drásticamente su saturación y sabíamos por los datos experimentales que todos los protones de la molécula se saturaban (Figura 5.3.3-A,B).

Si se combina la saturación calculada de las dos primeras orientaciones de la lista de R_e , parece que se describe mejor el patrón de STD-NMR experimental. Esta relación proporciona un indicio de la existencia de al menos dos modos de unión diferentes y explicaría tanto los altos valores de R_e como la poca saturación teórica en regiones de la dactilolida.

En un intento de mejorar el factor-R cristalográfico y refinar el mapa de densidad electrónica, se utilizaron las coordenadas de los cuatro complejos moleculares para optimizar las fases, calcular nuevamente las amplitudes (F_{calc}) y obtener los mapas de densidad electrónica (Figura 5.3.5). Las apariciones de nuevas densidades electrónicas fueron encontradas, tanto en β 1- como en β 2-tubulina, entorno a las coordenadas de la dactilolida, dispersadas por el resto de la cavidad y también fuera de ésta. A partir de las cuatro estructuras se obtuvieron mapas de diferencias negativas (enrejado rojo en la Figura 5.3.5) por la ausencia de ciertas porciones de la molécula en la estructura real, y regiones donde la densidad estaba desordenada o simplemente no existía. Por último, se observó que la densidad electrónica alrededor de la cadena lateral de la HIS229 estaba extendida, indicando la existencia de alguna posible interacción covalente (Figura 5.3.5-B).

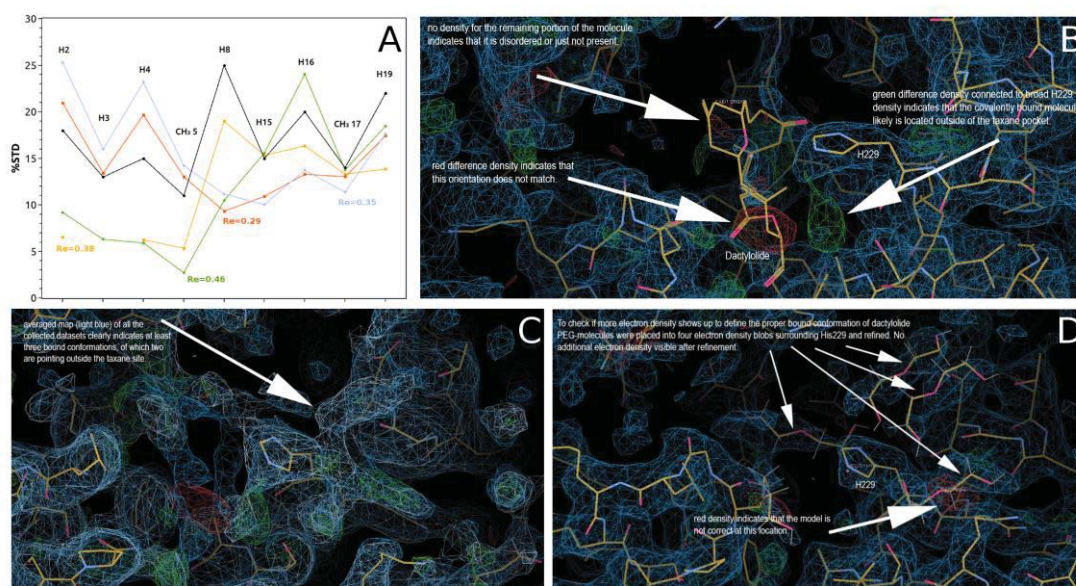


Figura 5.3.5. (A) Intensidades de STD-NMR del complejo de dactilolida- T_2R . Los datos experimentales se representan en negro, también se representan los datos calculados para el acoplamiento 1 en $\beta 1$ (rojo) y $\beta 2$ (azul claro), y el acoplamiento 2 en $\beta 1$ (amarillo) y $\beta 2$ (verde). (B) Refinamiento cristalográfico del acoplamiento 2 de la dactilolida y la $\beta 2$ -tubulina. (C) (D). Superposición de los mapas de densidad electrónica (F_{obs} , φ_{calc}) y los mapas de diferencias ($2F_{obs} - F_{calc}$, φ_{calc}). En todos los casos con una resolución de 2,5 Å. Los mapas de densidad electrónica están representados con una red azul. Los mapas de diferencia están representados en verde (contorno positivo, zona incluida en la estructura real y ausente en el modelo) y en rojo (contorno negativo, zona incluida en el modelo y ausente en la estructura real).

La información extraída a partir de los refinamientos cristalográficos fue ambigua. Sin importar la orientación de STD-NMR utilizada, mediante el nuevo refinamiento no se logró obtener la estructura del cristal. Este comportamiento es similar al descrito en los cálculos con el STD-MaPa. Así que ninguno de los cuatro modelos moleculares (2 con la DAC en $\beta 1$ -tubulina y 2 con la DAC en $\beta 2$ -tubulina) fue confiable, ya que las orientaciones extraídas a partir de los

datos de RMN indujeron la aparición de mapas electrónicos negativos o la ausencia de densidad electrónica.

Una explicación plausible a estos efectos puede basarse en la existencia de diversas orientaciones. Por eso, en un intento de clarificar esta hipótesis, los mapas de densidades recopilados hasta el momento se utilizaron para calcular un mapa promedio (Figura 5.3.5-C). Los resultados mostraron más de una posible orientación e incluso que alguna se encontrase localizada fuera de la cavidad. En un intento por enriquecer el mapa de densidades electrónicas, algunas moléculas de polietilenglicol (PEG) fueron distribuidas en las densidades electrónicas que se localizan alrededor de la HIS229 (Figura 5.3.5-D). Se refinó nuevamente el mapa de densidades, pero esta vez con las moléculas de PEG y no se encontraron densidades electrónicas adicionales. Así que, con toda la información obtenida a partir de los resultados de RMN y DRX, se puede decir que los datos cristalográficos concuerdan con los datos de RMN y las dos técnicas estarían describiendo la coexistencia de más de una orientación. Por consiguiente, se procedió a calcular el acoplamiento molecular con el STD-MaPa considerando múltiples modos de asociación.

5.3.5 Resolución de las estructuras de los múltiples modos de unión del complejo dactilolida- β -tubulina

La información extraída de los datos experimentales sobre la estructura del complejo dactilolida- β -tubulina ha permitido conocer que la dactilolida se orienta de múltiples formas cuando se asocia a la β -tubulina. Pero el número de orientaciones y la proporción se desconocen por completo; dado que las técnicas de RMN y DRX no han permitido resolver estas incógnitas. Por lo tanto, hay que buscar un método que permita calcular las múltiples asociaciones de la dactilolida y las relaciones entre las concentraciones de las uniones. En el apartado 5.2.7.1 se desarrolló una formulación general que

permite obtener las relaciones de concentraciones para describir múltiples modos de unión con el CORCEMA (EGC-CORCEMA) y que será utilizado para describir la unión de la DAC y el heterodímero de α,β -tubulina. Por tanto, si se desean resolver las ecuaciones generales de concentraciones del CORCEMA, se necesitan las concentraciones iniciales de la proteína y el ligando, la constante de asociación, el total de orientaciones del ligando y las relaciones de concentración entre los modos de unión.

En una primera aproximación, se debe considerar la situación más simple, en la cual el espectro de STD-NMR de la DAC describa solo dos orientaciones. Para este caso, se procede a optimizar la abundancia relativa entre las uniones considerando cinco relaciones de proporcionalidad. Si con esta aproximación, el patrón suma de STD-NMR no se ajusta los datos experimentales, se procederá a aumentar el número total de asociaciones, hasta que los factores R_e o R_t converjan. En los cálculos, las concentraciones iniciales del ligando y la proteína son las mismas que se utilizaron en el experimento de STD-NMR; así que para el complejo DAC- α,β -tubulina se utilizaron 300 μM de dactilolida y 10 μM de tubulina (ver apartado 4.5). En nuestro estudio, consideramos que la constante de asociación de la unión más abundante es igual a la constante de asociación aparente determinada en el estudio de desplazamiento de Flutax-2 ($K_a=0,6 \pm 0,1 \times 10^6 \text{ M}^{-1}$) (apartado 1.3). Finalmente, antes de utilizar el STD-MaPa en la resolución de la estructura de la DAC describiendo dos orientaciones, se calcularon las ecuaciones EGC-CORCEMA (Tabla 5.3.3).

Tabla 5.3.3. Parámetros y concentraciones resueltas a partir de las ecuaciones EGC-CORCEMA.

Proporción (k:i)	1:1	2:1	3:1	4:1	5:1
y_i^k	1	0,5	0,33	0,25	0,2
a_k	$4,00 \times 10^6$	$2,25 \times 10^6$	$1,78 \times 10^6$	$1,56 \times 10^6$	$1,44 \times 10^6$
b_k	-621	-466	-414	-389	-373
c_k	$3,0 \times 10^{-3}$	$3,0 \times 10^{-3}$	$3,0 \times 10^{-3}$	$3,0 \times 10^{-3}$	$3,0 \times 10^{-3}$
$[PL_k]$	$5,0 \times 10^{-6}$	$6,7 \times 10^{-6}$	$7,5 \times 10^{-6}$	$8,0 \times 10^{-6}$	$8,3 \times 10^{-6}$
$[PL_i]$	$5,0 \times 10^{-6}$	$3,3 \times 10^{-6}$	$2,5 \times 10^{-6}$	$2,0 \times 10^{-6}$	$1,7 \times 10^{-6}$
A_k	$3,33 \times 10^4$	$3,33 \times 10^4$	$3,33 \times 10^4$	$3,33 \times 10^4$	$3,33 \times 10^4$
B_k	-5,158	-6,873	-7,730	-8,244	-8,586
C_k	$2,0 \times 10^{-5}$	$3,8 \times 10^{-5}$	$4,9 \times 10^{-5}$	$5,6 \times 10^{-5}$	$6,1 \times 10^{-5}$
$c_{0,k}'(L)$	$1,5 \times 10^{-4}$	$2,0 \times 10^{-4}$	$2,3 \times 10^{-4}$	$2,4 \times 10^{-4}$	$2,5 \times 10^{-4}$
$c_{0,k}'(P)$	$5,0 \times 10^{-6}$	$6,7 \times 10^{-6}$	$7,5 \times 10^{-6}$	$8,0 \times 10^{-6}$	$8,3 \times 10^{-6}$
K_a^i	$1,00 \times 10^6$	$5,00 \times 10^5$	$3,33 \times 10^5$	$2,50 \times 10^5$	$2,00 \times 10^5$
A_i	$3,33 \times 10^4$	$1,67 \times 10^4$	$1,11 \times 10^4$	$8,33 \times 10^4$	$6,67 \times 10^4$
B_i	-5,158	-1,718	-0,859	-0,515	-0,343
C_i	$2,0 \times 10^{-5}$	$2,2 \times 10^{-6}$	$-4,2 \times 10^{-7}$	$-1,0 \times 10^{-6}$	$-1,1 \times 10^{-6}$
$c_{0,i}'(L)$	$1,5 \times 10^{-4}$	$1,0 \times 10^{-4}$	$7,8 \times 10^{-5}$	$6,4 \times 10^{-5}$	$5,5 \times 10^{-5}$
$c_{0,i}'(P)$	$5,0 \times 10^{-6}$	$3,4 \times 10^{-6}$	$2,6 \times 10^{-6}$	$2,1 \times 10^{-6}$	$1,8 \times 10^{-6}$

Las concentraciones están en unidades de molaridad. El modo de unión k-ésimo se considera como el predominante ($K_a^k = 10^6$). Las concentraciones iniciales se igualaron a $c_0(P) = 10^{-5}M$ y $c_0(L) = 3 \times 10^{-4}M$.

En la tabla Tabla 5.3.3 se muestran los datos calculados a partir de las ecuaciones de segundo orden. En dicha tabla se puede observar como la constante de asociación del segundo modo varía con la variación de la proporción entre los estados enlazados, disminuyendo con la menor presencia

del segundo modo. Este hecho ocurre porque en los cálculos se consideraron diferentes proporciones relativas entre las uniones, y la constante de asociación del modo más abundante se fijó en 10^6 M^{-1} . Además, las concentraciones iniciales ficticias de la proteína y ligando se calcularon para cada unión, lo que permitió resolver individualmente la estructura de cada orientación de la DAC con el STD-MaPa.

Hasta este momento, la última versión del programa STD-MaPa no incluye ni la aproximación de EGC-CORCEMA, ni una metodología automática que realice cada uno de los cálculos. En su lugar, se desarrolló un programa que fuera capaz de gestionar, organizar y evaluar todo el volumen de datos iniciales y finales del STD-MaPa (Figura 5.3.6). Este programa utilizaría el STD-MaPa e interactuaría directamente con los módulos encargados de evaluar las poses de unión. Entre las acciones que realizará la metodología, se encargará de sumar las intensidades de STD-NMR (ecuación 5.2.1) y de ejecutar el módulo de los factores R-NOE para crear las listas de evaluación a partir de los patrones suma.

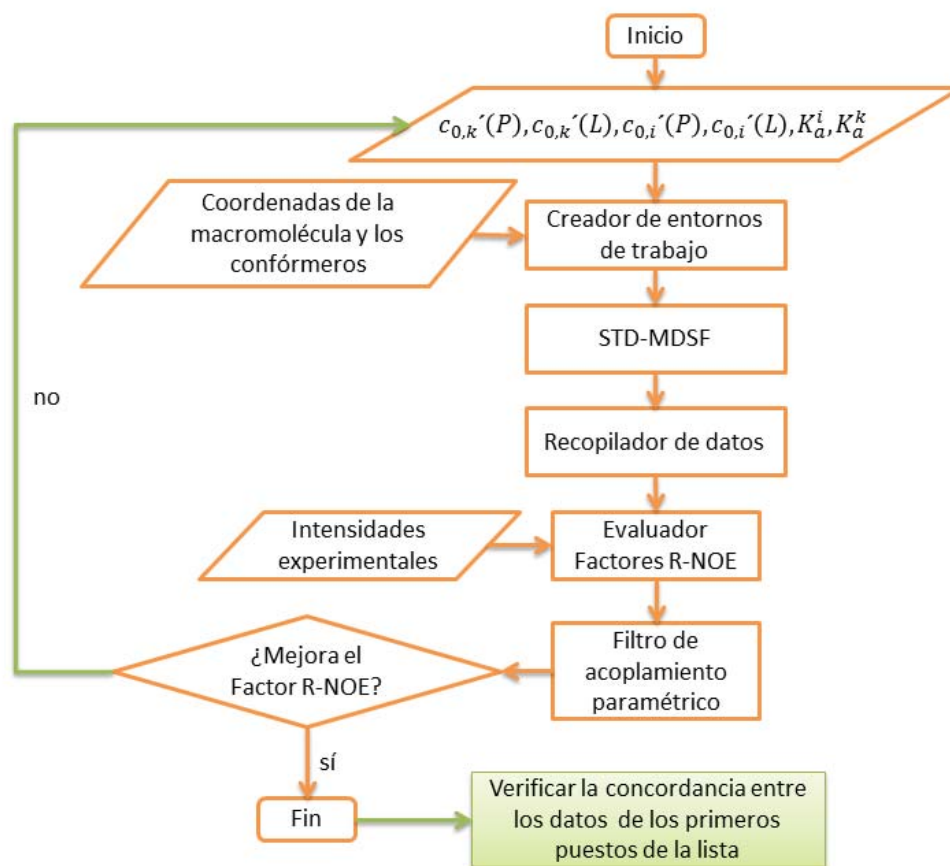


Figura 5.3.6. Diagrama de flujo del módulo externo al STD-MaPa encargado de resolver la estructura de los complejos moleculares con múltiples modos de unión. El STD-MaPa genera las coordenadas de los múltiples acoplamientos moleculares y optimiza los parámetros químico-físicos. En naranja se identifican las acciones realizadas por el módulo externo al STD-MaPa. En verde se identifican las acciones manuales y externas al programa.

El cálculo se realizó explorando la cavidad de β 1- y β 2-tubulina. Los patrones de intensidades resultantes se sumaron y los ajustes se evaluaron con los factores R_e , R_t y R_{tm} . Los resultados obtenidos mejoraron drásticamente comparados con los resultados previos. Los tres factores redujeron su valor drásticamente en los dos isotipos de tubulina. Además, con los dos modos de unión se llega a describir casi perfectamente el patrón de intensidades de STD-

NMR. Los factores escogieron entre las primeras cinco combinaciones el mismo par de orientaciones en β 1-tubulina y las mismas abundancias relativas (Figura 5.3.7). Además, los tres factores concordaron en la abundancia, dando un 83% de abundancia relativa al modo de unión mayoritario y un 17% para el modo de unión minoritario. Con esta proporción, la constante de asociación aparente del segundo modo de unión fue de $1 \times 10^5 \text{ M}^{-1}$. Además, R_e y R_t coincidieron en la asignación de la constante de velocidad de disociación y en los tiempos de correlación de la dactilolida libre, en los dos modos de unión en β 1-tubulina. Para las uniones con β 1-tubulina, los dos factores concordaron en establecer una constante de disociación aparente para la dactilolida de 10 s^{-1} y un tiempo de correlación de 0,5 ns. Por tanto, en las cinco primeras posiciones de las listas de los factores R-NOE se pueden encontrar todas las conformaciones de la dactilolida resueltas por TR-NOESY. El error relativo porcentual del factor R_e debido al cambio conformacional de las cinco primeras posiciones de su lista pudo ser calculado a partir del valor medio y la desviación estándar ($\overline{Re} = 0,121 \pm 0,004$). La variación obtenida representa solo el 3% del factor R_e , lo que llevaría a concluir que todos los confórmeros resueltos por TR-NOESY representan los experimentos de STD-NMR.

Mediante el mismo protocolo, se obtuvo el acoplamiento de la dactilolida a β 2-tubulina. En este caso se observó que las orientaciones obtenidas en β 2-tubulina eran similares a las resueltas en β 1-tubulina; pero las diferencias estructurales entre las dos cavidades produjeron un aumento del 17% en el factor R_e . Según este dato, las disposiciones espaciales que tienen los residuos en β 2-tubulina no son las mismas que las descritas por la estructura basada en el experimento de STD-NMR; aumentando 6 veces la desviación debida a los cambios conformacionales del ligando en β 2-tubulina.

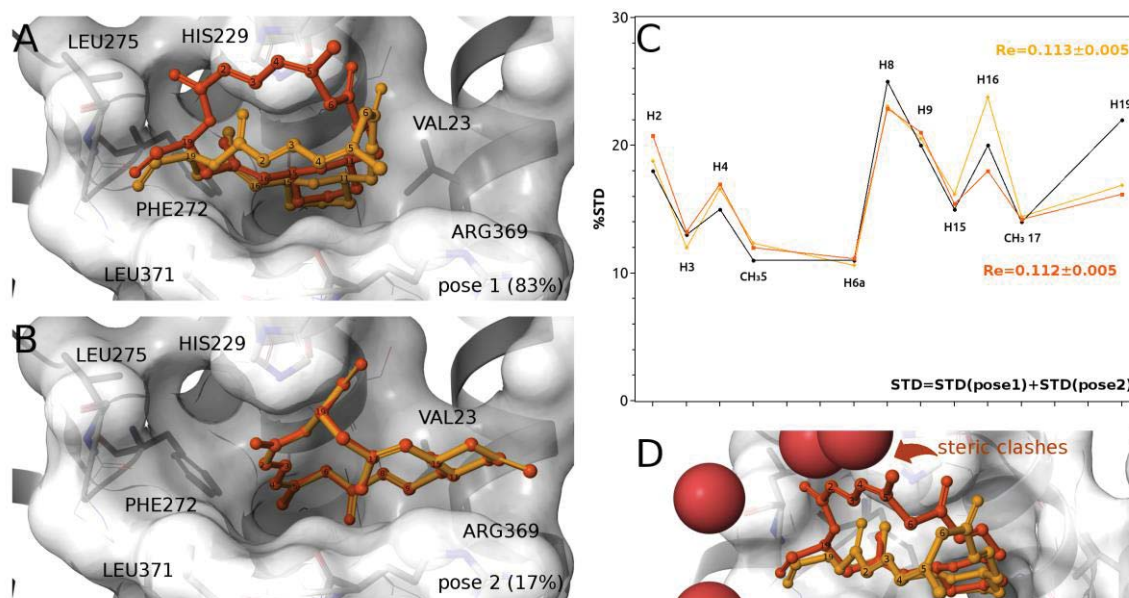


Figura 5.3.7. (A) Representación de los dos conformeros que minimizan el valor del factor R_e para el modo de unión mayoritario en $\beta 1$ -tubulina. (B) Representación de los dos conformeros que minimizan el valor del factor R_e para el modo de unión minoritario en $\beta 1$ -tubulina. (C) Distribución de intensidades de STD-NMR de los datos experimentales (negro) y los modos de unión resueltos por el STD-MaPa (rojo y naranja). Los colores de los datos corresponden con el par de estructuras representadas a la izquierda. (D) Representación de los modos de unión mayoritarios en presencia de los átomos de oxígenos de aguas resueltos para la estructura cristalográfica.

Por otro lado, al analizar la estructura cristalográfica (Figura 5.3.2-B), se encontró que la orientación del residuo HIS229 y la densidad electrónica alrededor del anillo de este residuo mostraban una mayor predisposición a un enlace covalente en $\beta 2$ -tubulina que en $\beta 1$ -tubulina. Por el contrario, la estructura más abundante y con menor R_e determinada a través del método híbrido (estructura localizada en $\beta 1$ -tubulina), presentó la HIS229 orientada hacia el centro del macrociclo de la dactilolida (Figura 5.3.7-A). Tal orientación no se corresponde con la forma covalente de la β -tubulina (Figura 5.3.2-B). Por

tanto, las orientaciones de los residuos de las cavidades de las dos β -tubulinas, en la estructura cristalográfica del complejo T₂R en presencia de dactilolida, describen dos procesos asociativos diferentes. Por las disposiciones espaciales de los residuos en β 2-tubulina el cristal describe la unión covalente de la dactilolida, mientras que en β 1-tubulina los residuos describen la asociación reversible.

A partir de estos resultados se puede decir que los datos de STD-NMR describen dos asociaciones diferentes e independientes entre la dactilolida y una β -tubulina con la disposición espacial descrita por la β 1-tubulina del cristal. A partir del análisis realizado a las estructuras obtenidas por el método híbrido, podemos decir que la orientación más abundante está enterrada completamente en la cavidad, como describían los datos de RMN. El anillo 4-metilen-tetrahidropirano está ocluido entre la hélice H7, la hélice H1 y el lazo B9-B10; en el interior de un bolsillo hidrofóbico compuesto por la PHE272, GLU27, VAL23, SER236 y PRO360. Además, por la cercanía, la DAC tiene la posibilidad de formar un puente de hidrógeno a través del grupo aldehído con la cadena lateral de la LEU275. A su vez, el grupo butadieno queda orientado hacia el exterior del sitio, localizado en la región más alejada de la proteína, lo que le confiere una mayor libertad conformacional y explica que los datos de RMN no distingan entre los confórmeros de TR-NOESY (Figura 5.3.7-C).

Por otra parte, la orientación menos abundante se encuentra solo parcialmente en el sitio de unión. La mayor parte de la molécula está dirigida hacia el exterior oriental del sitio (Figura 5.3.7-B). Esta estructura 3D logra explicar parte de la densidad electrónica situada en la región más oriental entorno a la HIS229 (Figura 5.3.5-C). En esta orientación el grupo butileno se encuentra dentro del sitio de unión, el grupo aldehído tiene la capacidad de formar un puente de hidrógeno con la HIS229 y el CH₃-5 queda situado frente a la PHE272.

Al considerar los dos modos de unión, el STD-NMR suma describiría perfectamente los datos de STD-NMR experimentales (Figura 5.3.7-C). Solo la intensidad de STD-NMR del protón H19 es menor a la descrita por el

experimento de STD-NMR. El H19 pertenece al aldehído de la cadena lateral de la dactilolida (Figura 5.3.2 o Figura 5.3.3) y se encuentra orientado hacia el lazo-M en el modo de unión predominante. La ausencia del lazo-M en el modelo resuelto a partir del método híbrido reduce ligeramente la intensidad del STD-NMR calculado. Sin embargo, el resto de los protones de la molécula no parecen estar afectados por la ausencia de la estructura del lazo-M.

En la comparación de los datos de STD-NMR calculados y experimentales se observa como la intensidad del protón H16 cambia drásticamente entre confórmeros de DAC (Figura 5.3.7-C). H16 está enterrado en la cavidad de β -tubulina cuando la DAC se encuentra en la orientación mayoritaria. Así que, cuando la estructura del macrociclo pasa de la posición plana (DAC de color rojo en Figura 5.3.7-A,B) a la plegada (DAC de color naranja en Figura 5.3.7-A,B), las coordenadas de H16 se modifican ligeramente, reduciendo la distancia con la PRO360 de la proteína. Este desplazamiento que sufre H16 entre un confórmero y otro se refleja en el aumento y la disminución de la intensidad de la saturación del protón en los modelos obtenidos con el método híbrido (Figura 5.3.7-C).

Los cálculos de STD-NMR se realizan a partir de estructuras rígidas, pero la realidad es diferente. La información extraída por los experimentos de STD-NMR proviene de todas las orientaciones y conformaciones accesibles del sistema, por lo tanto la intensidad experimental de H16 es un valor medio. Como se puede apreciar en la Figura 5.3.7-C, el valor experimental de H16 se encuentra entre los valores de los dos modelos moleculares. En la orientación predominante, la mayor parte de la estructura del ligando está muy bien definida al encontrarse completamente ocluida en la cavidad de β -tubulina. Sin embargo, la región de la DAC más expuesta no interacciona con la proteína, lo que aporta cierta flexibilidad a la estructura y favorece los cambios de la conformación en esta zona de la molécula. En estos casos, el comportamiento de las intensidades en el protón H16 es un indicio de la descripción, por parte del experimento de STD-NMR, de múltiples conformaciones del ligando.

La variación en el patrón de intensidades y el factor R_e no permiten reconocer la conformación que mejor describe los datos de STD-NMR. Igualmente, esta aproximación carece de sentido frente a las diferencias en el R_e , debido a la información dinámica implícita en las técnicas de RMN. Sin embargo, la densidad electrónica del cristal describe la posición de dos moléculas de agua a la salida del sitio de $\beta 2$ -tubulina (Figura 5.3.7-D). Los oxígenos son un indicio de que la DAC en el cristal no adquiere la conformación alargada (DAC en rojo en la Figura 5.3.7-A,D), dado que produciría choques estéricos. Por lo tanto, la estructura mayoritaria resuelta por el método híbrido, tiene al complejo molecular DAC- T_2R con el ligando en la conformación plegada (DAC en naranja en la Figura 5.3.7-A,D), ya que sus coordenadas describen mejor los datos de RMN (TR-NOESY y STD-NMR) y los datos cristalográficos. Así que, a partir del método híbrido RMN-DRX, se resolvió la estructura del complejo de dactilolida- T_2R describiendo una unión doble con $\beta 1$ -tubulina.

Discusión

6 Discusión

6.1 Nuevos factores R-NOE

El primer objetivo del trabajo estuvo dirigido al diseño de nuevos factores R-NOE (R_t , R_{tm} , R_{twm} , R_{ms} y R_{msw}) para la evaluación de estructuras ligando-receptor a partir de datos de RMN. La resolución estructural de complejos formados por macromoléculas biológicas y moléculas pequeñas es importante cuando se desea entender el funcionamiento de los mecanismos celulares. Entre las herramientas que se utilizan para obtener la estructura de los complejos se encuentra la diferencia transferida por saturación (STD-NMR). En capítulos anteriores se ha comentado que el STD-NMR es un método de resonancia magnética nuclear que, con un procedimiento experimental sencillo, permite extraer el patrón de NOEs entre el ligando y la macromolécula.

Esta técnica es de nuestro interés porque a partir de la información recogida en los experimentos de STD-NMR se resuelve el acoplamiento de complejos moleculares. En el procedimiento de resolución estructural se utiliza el factor R-NOE como función de coste para evaluar la estructura modelada, mediante la comparación de la saturación protónica de la estructura y los datos experimentales. Este factor proviene de la función de error residual de difracción de rayos-X, y se adaptó para aplicarla en la resolución estructural de complejos moleculares basada en datos de STD-NMR (factor R_e). En la adaptación, el término ($W_{exp,k}$) fue añadido porque permite que el factor R_e sea sensible tanto a la variación de los valores de STD-NMR pequeños como a la desviación en valores grandes de STD-NMR.

El acoplamiento entre espines depende de la sexta potencia del inverso de la distancia (r^{-6}). Por lo tanto, cuando la distancia entre dos protones (r) es pequeña, un ligero cambio en su valor modifica sustancialmente la relajación de la saturación entre los dos protones. Por el contrario, en presencia de

valores de distancia elevados (entre 5 y 8 Å), los cambios en las coordenadas de la estructura no inducen modificaciones substanciales elevadas en los valores de la saturación. Los dos efectos se trasladan al factor R-NOE y, si no se tienen en cuenta, provocan que el factor R-NOE describa un comportamiento errático.

Aunque la adición de $W_{\text{exp},k}$ fue acertada, el término sobrestima los datos que generalmente provienen de señales espectrales con menor relación señal/ruido. Sobrestimar las señales peor resueltas del espectro puede incrementar la incertidumbre del factor R-NOE. Sin embargo, el término $W_{\text{exp},k}$ es importante para la correcta evaluación de las estructuras con el factor R-NOE; por lo que consideramos el término $W_{\text{cal},k}$ como sustituto de $W_{\text{exp},k}$. El nuevo término actúa sobre las intensidades del modelo molecular teórico ($S_{\text{cal},k}$) y, con él, se modula la saturación de los protones que en el modelo están ubicados muy cerca o muy lejos de la superficie de la proteína.

Tanto $W_{\text{exp},k}$ como $W_{\text{cal},k}$ modulan los cambios extremos en la saturación mediante la ponderación del término resta ($S_{\text{exp},k} - S_{\text{cal},k}$) del factor R-NOE (ecuaciones 1.2.10 y 5.1.1). La diferencia entre ambos términos radica en la referencia que utilizan; ya que el conjunto de $W_{\text{exp},k}$ pondera los términos restas de la ecuación del factor R-NOE considerando la estructura real (estructura que mejor se ajusta al patrón de STD-NMR), y el conjunto de $W_{\text{cal},k}$ lo hace considerando cada una de las estructuras que se evalúan. Por tanto, al evaluar el modelo molecular que mejor se ajusta al patrón de STD-NMR (modelo real), los dos términos tienden a ser iguales ($W_{\text{exp},k} \approx W_{\text{cal},k}$).

$W_{\text{cal},k}$ no depende de los datos experimentales, por lo que se espera que el conjunto de $W_{\text{cal},k}$ no aumente la incertidumbre de la evaluación al no sobreestimar las señales peor resueltas del espectro. Para realizar un análisis más profundo sobre la incertidumbre que añade el término $W_k^{\text{cal o exp}}$, en el anexo 9.1 se resolvieron las ecuaciones diferenciales del factor R_e y R_t (ecuaciones 9.1.5 y 9.1.6). Las dos ecuaciones fueron resueltas para poder propagar los errores al factor R-NOE a partir de los términos independientes

$S_{exp,k}$ y $S_{cal,k}$ y obtener las incertidumbres correspondientes al factor R_e (en presencia de $W_{exp,k}$) y R_t (en presencia de $W_{cal,k}$) (ecuaciones 9.1.1 y 9.1.2).

Si se analizan las ecuaciones diferenciales, se puede observar que dependen de W_k , $S_{exp,k}$ y $S_{cal,k}$. Además, las dos tienen dos términos en común (primer y segundo término de las ecuaciones diferenciales), y solo $\delta R_e / \delta S_{exp,k}$ se compone de cuatro términos. Según las ecuaciones 9.1.5 ($\delta R_e / \delta S_{exp,k}$) y 9.1.6 ($\delta R_e / \delta S_{cal,k}$), los cuatro términos mantienen una relación directa con $W_{exp,k}$ y $W_{cal,k}$. Por lo tanto, cuando la relación $S_{exp,k} / dS_{exp,k}$ disminuye, $W_{exp,k}$ siempre introduce un aumento en los cuatro términos de la ecuación 9.1.5. Pero, aunque esto ocurra, el error de R_e no aumenta porque los cuatro términos compensan el aumento introducido por $W_{exp,k}$.

Para entender el aporte de cada término sobre el error y conocer en más detalle la relación entre $W^{exp \text{ o } cal}_k$ y el factor R-NOE, se estudió un sistema hipotético compuesto por un ligando con 24 protones. La saturación experimental para todos los protones se consideró igual a 12 ($S_{exp}=12$) y para la saturación teórica se generaron 24 valores en orden creciente que van desde $S_{cal}=1$ hasta $S_{cal}=24$ (Figura 6.1.1). En este hipotético caso, como todos los protones reciben la misma saturación experimental, los términos 2 y 4 son constantes. Para facilitar el análisis, la suma del primer y segundo término de la ecuación diferencial de R_e será $f(x)$ y la suma de los términos 3 y 4 será $g(x)$.

Tanto en la ecuación diferencial 9.1.5 como en la Figura 6.1.1 se puede observar como los términos $f(x)$ y $g(x)$ dependen de la diferencia entre la saturación experimental y la calculada. Por otro lado, los términos 2 y 4, al ser siempre negativos, desplazan la función $f(x)$ hacia regiones más negativas del eje de ordenadas y la función $g(x)$ hacia regiones más positivas. Este comportamiento compensa los posibles aumentos en $W_{exp,k}$ y provoca que los errores de R_e y R_t sean similares.

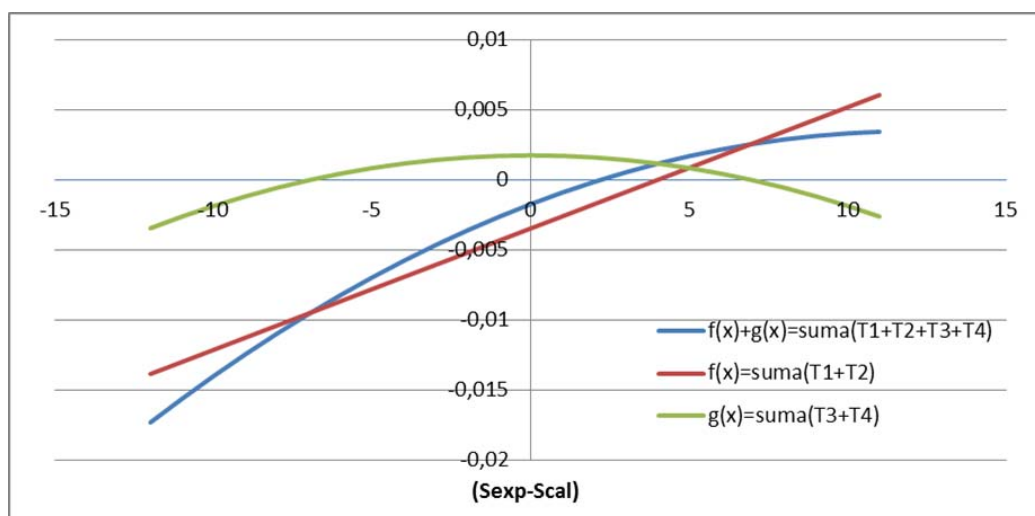


Figura 6.1.1 Variación de los cuatro términos de $\delta R_e/\delta S_e$ en función de la diferencia entre el valor experimental y el teórico. En el cálculo se consideró un ligando hipotético con 24 protones con la misma saturación experimental ($S_{\text{exp}}=12$). Los valores de la saturación teórica (S_{cal}) se fijaron con valores crecientes de saturación que van desde $S_{\text{cal}}=1$ hasta $S_{\text{cal}}=24$.

El cálculo de los errores de R_e y R_t tiene ciertas ventajas adicionales que repercuten sobre el proceso predictivo implementado en el STD-MaPa. Primero, los errores permiten conocer la significación estadística entre los modelos teóricos que se evalúan con el factor R-NOE. Además, el error en ambos casos depende de la diferencia entre la saturación experimental y calculada; de manera que el error se reduce en presencia de la estructural real, y aumenta cuando las diferencias estructurales entre el modelo real y el calculado son elevadas. Con la introducción de esta nueva información, se podrán identificar el modelo o los modelos que mejor ajustan el patrón de STD-NMR experimental de un modo más cuantitativo.

Contrario a lo que se esperaba, la incertidumbre que introduce $W_{\text{exp},k}$ es similar a la que introduce $W_{\text{cal},k}$, así que los dos términos son igualmente válidos en el análisis de los modelos moleculares. Al tener incertidumbres parecidas, la capacidad de reconocimiento de los factores R_e y R_t es similar; como se puede constatar en los estudios realizados en los apartados 5.2.5 y 5.2.6. En estos

estudios, los dos factores llegaron a reconocer el 100% de las estructuras buscadas.

En la construcción de los factores R_{tm} y R_{twm} se añadió $W_{cal,k}$, porque se tomaron como referencias las intensidades de la estructura calculada, y así se logró una mayor adaptación a los modelos moleculares generados con el STD-MaPa. Los factores R_{tm} , R_{twm} , R_{ms} y R_{msw} se diseñaron adicionalmente para rectificar las desviaciones producidas por la dependencia del STD-NMR debida a los parámetros k_{off} , k_{on} , τ_c^L y τ_c^P . Las cuatro funciones tienen el término ΔM y ΔM_w para rectificar la desviación media del patrón de saturación.

Mediante la introducción de los términos ΔM y ΔM_w se llegó a reducir entre un 40 y un 50% la sensibilidad que presentan los factores R_e y R_t por los parámetros químico-físicos (Figura 5.1.5). Con esta reducción logramos funciones más robustas, capaces de rectificar los efectos producidos por los cambios en el tamaño de la red dipolar (τ_c), la velocidad de transferencia de la saturación hacia la solución (k_{off}); así como de compensar la pérdida de la transferencia de la saturación al explorar constantes de asociación con valores menores a 10^4 M^{-1} . Sin embargo, entre los dos términos, ΔM subestima los protones menos saturados cuando las diferencias entre las intensidades teóricas difieren significativamente en un mismo ligando. Por el contrario, ΔM_w logra rectificar estos efectos, mediante la ponderación de la media del patrón de saturación con el término de peso $W_{cal,k}$. Sin embargo, los sistemas moleculares utilizados en nuestros estudios presentan patrones de saturación homogéneos, y debido a las pocas diferencias en la saturación entre los protones del mismo ligando, nuestros resultados no mostraron diferencias significativas entre ΔM y ΔM_w (Figura 5.1.3).

Estos dos factores influyen notoriamente en el comportamiento del factor R-NOE y su influencia sobre ellos es sólida. Esta afirmación se puede argumentar analizando los factores R_{ms} y R_{msw} . En estos dos factores, el efecto de ΔM y ΔM_w se tiene en cuenta solo un 50% y, además, la desviación de la saturación

teórica se penaliza con el término $N_i\Delta M^2$ o $N_i\Delta M_w^2$ (ecuaciones 5.1.6 y 5.1.7). Los resultados de la Tabla 5.1.5 muestran como ΔM y ΔM_w , a expensas de $N_i\Delta M^2$ y $N_i\Delta M_w^2$ (R_{ms} y R_{msw}), reducen la dependencia paramétrica del factor R-NOE casi igual que en ausencia de los términos de penalización (R_{tm} y R_{twm}). En la misma tabla, al comparar los errores relativos entre el grupo de factores R-NOE con ΔM o ΔM_w y el grupo que no tiene ninguno de los dos (R_e y R_t), se observa como los resultados con el grupo de ΔM y ΔM_w son poco precisos en las dos primeras posiciones de las listas generada por los factores y selectivos si se consideran las cinco primeras posiciones. En este estudio, se cuantificó la cantidad de aciertos ubicados en las cinco primeras posiciones de cada una de las listas generadas con los factores R-NOE, en la búsqueda de la estructura real de tres sistemas moleculares entre 300 opciones. Los aciertos de los tres sistemas moleculares estudiados fueron promediados y las desviaciones estándar calculadas.

El error relativo fue calculado a partir de los aciertos de los dos primeros puestos de cada lista. En este caso, R_t presentó el menor error con un 12,5%, seguido de R_e con un 40%, R_{twm} y R_{msw} con un 100% y finalmente R_{tm} y R_{ms} con un 170 y 150% respectivamente. Los resultados de este estudio muestran como la substitución de $W_{exp,k}$ por $W_{cal,k}$ induce una mejora en la precisión del factor R-NOE, al adaptarse a la estructura modelada y no a la real. Mientras, la adición de ΔM o ΔM_w afectó drásticamente los resultados de los factores R-NOE, reduciendo las capacidades de R_{tm} , R_{twm} , R_{ms} y R_{msw} de discernir la estructura real.

Sin embargo, este hecho solo ocurre cuando consideramos las dos primeras posiciones de la lista. Por el contrario, cuando se analizan los resultados de las primeras cinco posiciones de las listas de los factores R_{tm} , R_{twm} , R_{ms} y R_{msw} , se llegan a identificar las estructuras reales el 100% de la veces con independencia de la parametrización utilizada. El resultado nos aporta una importante ventaja si se desea determinar la estructura de un complejo poco estudiado, ya que, con la aplicación de los factores R_{tm} , R_{twm} , R_{ms} y R_{msw} , no se

necesita información previa para determinar la orientación y ubicación del ligando en el sitio de unión de la macromolécula. La selectividad de estos cuatro factores en la identificación de la estructura real confirmó que la adición de los términos de peso (ΔM o ΔM_w) reducen los efectos de los parámetros químico-físicos sin reducir la sensibilidad de los factores R-NOE a los cambios en las coordenadas de los modelos moleculares.

Para comprobar que los términos añadidos al factor R-NOE no cambian su sensibilidad al cambio de coordenadas, se analizó la exploración torsional realizada en los complejos moleculares de salacinol, kifunesina y epotilona-A (Figura 5.1.7). Al comparar las variaciones que describieron los factores R_t , R_{tm} , R_{twm} , R_{ms} y R_{msw} con el factor R_e en las imágenes 2D se aprecia que las modificaciones realizadas sobre el factor R-NOE no modifican sensiblemente su respuesta al cambio de coordenadas, aunque sí afecta la profundidad de los máximos y mínimos de las superficies (Figura 5.1.7). Los términos ΔM o ΔM_w son los causantes de reducir la profundidad de los extremos de las superficies (máximos y mínimos) de los factores R_{tm} , R_{twm} , R_{ms} y R_{msw} , y tal reducción provoca que los cuatro factores (R_{tm} , R_{twm} , R_{ms} y R_{msw}) necesiten cinco oportunidades para llegar a discernir la estructura real (Tabla 5.1.5). Por el contrario, las superficies de R_e y R_t tienen extremos más agudos, lo que implica una mejor respuesta frente a cambios más pequeños en las coordenadas; siempre que se optimicen los parámetros o se conozca la parametrización adecuada. Así que, con ellos, se podrán optimizar mejor las estructuras e identificar entre varias conformaciones con un bajo valor de RMSD. Ése fue el caso de la epotilona-A, ya que en la resolución de su estructura unida a la β -tubulina, los factores R_e y R_t diferenciaron dos conformaciones con un RMSD entre las dos de solo 0,8 Å.

Como resultado del trabajo se consiguieron diseñar cinco factores R-NOE, que junto con el factor R_e permiten reconocer las coordenadas que mejor describen el patrón de STD-NMR. Con ellas, se llegaron a reconocer las estructuras reales de diversos complejos moleculares. Por último, se logró determinar la

estructura del complejo dactilolida-tubulina mediante la aplicación del factor R_e y R_t , permitiendo su caracterización estructural.

6.2 STD-MaPa es capaz de generar estructuras válidas a pesar de que la aproximación CORCEMA no contemple los efectos de reasociación y de la relajación T_1

Nuestro segundo objetivo fue la implementación de una metodología automática y autoconsistente que permite obtener las coordenadas de un complejo receptor-ligando a partir de datos de STD-NMR. Con este fin, se diseñó y creó el programa STD-MaPa. Este programa ha sido capaz de resolver diversas estructuras, incluso aunque no se hayan contemplados diversos efectos experimentales que modifican el patrón de STD-NMR.

Hasta estos momentos, en los cálculos de las intensidades de STD-NMR no se han tenido en cuenta todos los efectos que están modificando la magnetización de los protones en los experimentos. En un experimento existe un grupo de efectos que no están relacionados con el efecto NOE y que afecta la saturación de los protones del ligando de manera diferenciada. Estos efectos causan una desviación diferenciada en cada protón del ligando sobre las intensidades de los STD-NMR experimentales, que deforman el patrón de STD-NMR. Entre los factores que más afectan a las desviaciones se encuentran las diferentes capacidades de los protones del ligando para acumular saturación debido a la relajación longitudinal (T_1) y el proceso de reasociación de las moléculas de ligando previamente saturadas. En el caso de la relajación longitudinal (T_1), ésta se compone de un conjunto de procesos mediante los cuales se recupera la magnetización en equilibrio térmico. Mientras que, en la reasociación, moléculas de ligando previamente saturadas y con menos capacidad para captar la saturación transferida desde la macromolécula, se unen más de una vez al sitio de unión. La aproximación CORCEMA no incluye estos procesos;

por lo que, si nuestro algoritmo no los considera o no es capaz de rectificarlos, éste puede dar como resultado soluciones estructurales erróneas.

La relajación longitudinal T_1 es uno de los procesos que inducen ciertas desviaciones en los datos de STD-NMR, incrementándose su efecto a tiempos de saturación largos. Con estos tiempos de saturación (típicamente $t_s > 5$ s) se alcanza un máximo en los valores de Diferencia de Transferencia de Saturación, porque las velocidades de saturación y de relajación se igualan. Cuando esto ocurre, por medio de evidencias experimentales se ha encontrado que el patrón de intensidades se encuentra controlado por la relajación T_1 (Yan, Kline et al. 2003). Sin embargo, cuando los tiempos de saturación (t_s) son mayores (típicamente se satura la muestra entre 1 y 2 s) que el menor T_1 ($\sim 0,5$ -2 s), ya se empiezan a encontrar desviaciones en la saturación, las cuales son dependientes del valor T_1 de cada protón (Yan, Kline et al. 2003). Así, por ejemplo, la presencia de protones aromáticos (característicamente T_1 largo) junto con protones de un grupo metilo (T_1 corto) en un mismo ligando y con un tiempo de saturación largo (> 5 s), podría provocar una sobrestimación de la distancia entre el grupo metilo y la superficie de la molécula diana en la estimación estructural, debida a la reducción de la saturación sobre el grupo químico.

Por otra parte se tienen además los efectos de reasociación de moléculas de ligando previamente saturadas. En estos casos, el acceso de una molécula previamente saturada, y probablemente con menos capacidad de recibir saturación por parte del receptor, impide que otra molécula del mismo compuesto pueda entrar al sitio de la diana molecular. Este hecho repercute sobre el mecanismo de transferencia de saturación a través de las moléculas de ligando hacia la solución, provocando una disminución en la acumulación de la saturación de los ligandos libres. El efecto producido por la reasociación se agudiza cuando el tiempo de residencia del ligando en la solución (τ_{res}^l) es menor que T_1 . Cuando se cumple esta condición, la velocidad de reasociación es más rápida que los procesos de relajación, y las moléculas de ligando

previamente saturadas no tendrán tiempo de relajar su saturación antes de unirse nuevamente a la macromolécula. Este hecho contribuirá a la reducción de las intensidades de STD-NMR (Angulo, Enríquez-Navas et al. 2010).

Los dos efectos pueden evitarse o al menos reducirse, si se toman ciertas precauciones en el experimento (Meyer and Peters 2003, Yan, Kline et al. 2003). Entre las precauciones que se pueden tomar, la más efectiva de todas radica en la reducción del tiempo de saturación a valores menores que la mitad del menor T1 o τ_{res}^1 , dependiendo de la escala de tiempos de ambos procesos. Pero esta aproximación reduce drásticamente la sensibilidad del STD-NMR, ya que con cortos períodos de saturación selectiva (Figura 1.2.9 del apartado 1.2.4) se obtendría una macromolécula insuficientemente saturada (Yan, Kline et al. 2003). Una solución alternativa fue planteada por Mayer y James. Estos autores propusieron una curva de crecimiento monoexponencial con los STD-NMR experimentales (ordenadas) y diferentes tiempo de saturación (abscisas) (Mayer and James 2004).

$$STD_t = STD_{max}(1 - e^{-k_{sat}t_{sat}}) \quad 6.2.1$$

En la ecuación, STD_t es la intensidad de STD-NMR de un protón a un tiempo de saturación t_{sat} , STD_{max} es la máxima intensidad que el protón alcanzará con largos tiempos de saturación y k_{sat} es la constante de velocidad de saturación observada. A partir de la ecuación 6.2.1, se obtuvo la derivada y con ella hallaron la pendiente de la curva a tiempo de saturación cero.

$$dSTD_0 = STD_{max}k_{sat} \quad 6.2.2$$

Su propuesta consistía en describir la unión de los ligandos a partir de la pendiente inicial ($dSTD_0$); ya que se considera que $dSTD_0$ es un reflejo más

preciso de la proximidad entre los protones del ligando y el receptor. La pendiente inicial se encarga de describir el tiempo ($1/k_{\text{sat}}$) necesario para llegar a obtener el valor de saturación STD_{max} si se mantuviera la misma velocidad de saturación desde el inicio. Dado que con tiempos de saturación muy pequeños los mecanismos de relajación T1 y de reasociación no han iniciado, el dSTD_0 no se encuentra influenciado por sus efectos. Sin embargo, esta aproximación tiene como inconveniente que incrementa el tiempo del experimento, lo que impide su aplicación en el estudio de macromoléculas que se degraden a la temperatura de medición (típicamente, la temperatura ambiente).

Como ya se ha comentado, los efectos que provocan las desviaciones en los patrones de STD-NMR no fueron considerados en la formulación del CORCEMA y, como nuestra metodología utiliza esta aproximación para calcular los patrones de STD-NMR en los modelos moleculares, el procedimiento implementado en el STD-MaPa tampoco tiene en cuenta la presencia de estos artefactos en los datos de STD-NMR. Sin embargo, el protocolo implementado en el STD-MaPa resolvió las estructuras de cinco complejos moleculares, a partir de sus patrones de STD-NMR y sin utilizar las pendientes de saturación inicial a tiempo cero (dSTD_0). Además, en los cinco casos, no se utilizaron tiempos de saturación cortos; en ellos se saturó selectivamente la proteína durante ~ 2 s ($t_{\text{sat}} \approx 2$ s). Además, en todos los casos se obtuvo un buen ajuste en las intensidades de STD-NMR mediante la minimización del factor R_e y R_t (Figura 5.2.11). Así que, según los resultados obtenidos en el apartado 5.2.6, las desviaciones en la saturación no afectaron la capacidad de la metodología del STD-MaPa para resolver las estructuras de los complejos moleculares; ya que, con el programa se llegó a corregir las desviaciones en la saturación.

Para explicar este fenómeno fueron planteadas dos hipótesis. En la primera se considera que, si las intensidades de STD-NMR presentan desviaciones diferenciadas, el STD-MaPa homogenizaría todas las desviaciones del experimento en la evaluación de los STD-NMR calculados mediante los

factores R-NOE. Para comprobar la primera hipótesis, se analizaron las contribuciones de los términos $W_{\text{exp},k}$, $W_{\text{cal},k}$, ΔM y ΔM_w . Las contribuciones de $N_i \Delta M^2$ y $N_i \Delta M_w^2$ no fueron analizadas porque los dos términos influyen por igual en todo el patrón de STD-NMR (ecuaciones 5.1.6 y 5.1.7).

El primer término analizado fue $W_{\text{exp},k}$. Si se recuerda, este término sólo sobrestima las saturaciones menos intensas y este hecho no corrige las desviaciones. Por el contrario, el término podría estar sobrestimando las señales más afectadas por T1 y la reasociación. Por otra lado se encuentra el término $W_{\text{cal},k}$, que dada su naturaleza, tampoco elimina las desviaciones de la saturación. Sin embargo, ambos términos ($W_{\text{cal},k}$ y $W_{\text{exp},k}$) podrían rectificar las desviaciones en la saturación de los protones, siempre que, por cierta casuística, coincidiesen las magnitudes de los efectos de la relajación T1 y los efectos de la reasociación sobre cada protón, con la diferencia entre la saturación calculada y experimental, situación que es improbable.

Por otro lado, los términos ΔM y ΔM_w sí son capaces de rectificar las desviaciones en el patrón de STD-NMR. Con estos términos se consiguió que la desviación diferenciada de la saturación en cada protón, debida a artefactos experimentales, se redistribuya entre todas las señales (Figura 6.2.1). A través de la substracción de las diferencias entre el valor medio de las señales experimentales y el valor medio de los datos de STD-NMR calculados, se distribuyen, como se puede apreciar en la Figura 6.2.1, las desviaciones mediante el ajuste de las intensidades del patrón de STD-NMR teórico. En los cuatro ejemplos que se muestran en la imagen se puede observar cómo, al aplicar los términos ΔM y ΔM_w , se traslada la diferencia entre el par de señales marcadas con un asterisco hacia el segundo par de señales. Este simple artefacto matemático favoreció que los factores R_{tm} , R_{twm} , R_{ms} y R_{msw} no se vieran afectados por la relajación T1 o la reasociación en los estudios realizados en el apartado 5.1.7.

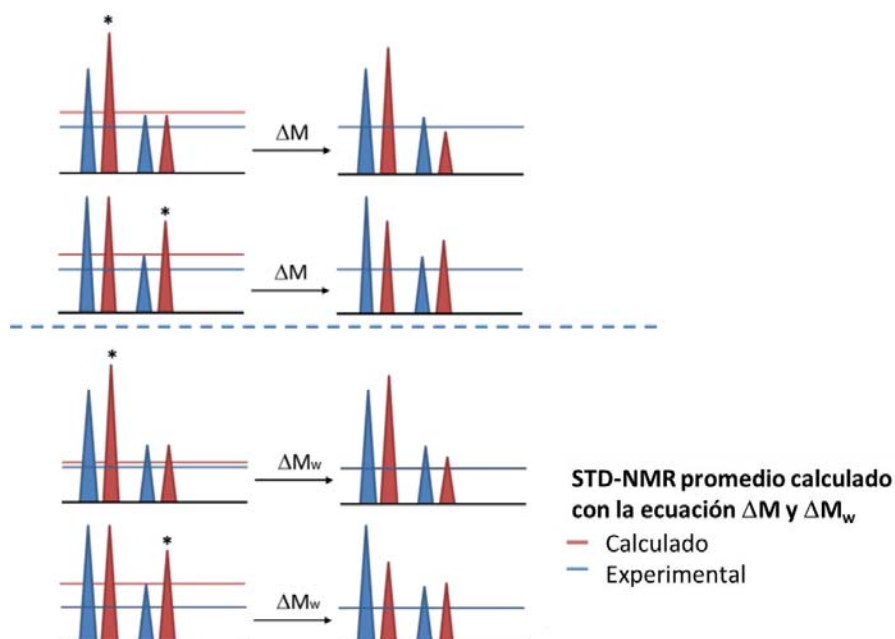


Figura 6.2.1 Representación esquemática de la redistribución de la saturación efectuada mediante los términos ΔM y ΔM_w en presencia de desviaciones en la saturación experimental debida a artefactos experimentales. Las líneas representan la media aritmética y ponderada de las intensidades.

Independientemente de la anterior hipótesis, todavía no se han expuesto las causas que fundamentarán la correcta descripción de los patrones de STD-NMR experimentales (Figura 5.2.11). La anterior hipótesis sólo es aplicable al reconocimiento molecular con el STD-MaPa, y este hecho no necesariamente asegura que los patrones experimentales y teóricos se ajusten correctamente.

En la búsqueda de evidencias que expliquen el comportamiento del STD-MaPa frente a las desviaciones en el patrón de saturación, se analizó el comportamiento del STD-NMR en función de diferentes parámetros (apartado 5.1.3). En el estudio de la Figura 5.1.2 del apartado 5.1.3 se calcularon los STD-NMR promedios de tres complejos moleculares a partir de los patrones de STD-NMR teóricos obtenidos con el CORCEMA-ST, en función de la variación de la constante de asociación (Figura 5.1.2, segunda fila) y del tiempo de

correlación de las proteínas (Figura 5.1.2, primera fila). En la representación, fueron tomados como referencias los valores promedios de cada patrón experimental. En la imagen se visualiza con claridad cómo las diferencias entre los datos experimentales y teóricos (la saturación promedio y la desviación estándar) se reducen en los tres ligandos para una combinación específica de parámetros químico-físicos; pudiendo concluirse a partir de estos resultados que, en los tres ligandos (kifunesina en la primera columna, salacinol en la segunda columna y epotilona-A en la tercera columna), se llegó a reproducir tanto el comportamiento medio del patrón de STD-NMR experimental como sus desviaciones.

Para explicar estos resultados, se necesitarán conocer los efectos que producen los mecanismos de relajación T1 y de reasociación sobre el STD-NMR; y éstos se pueden saber al analizar matemáticamente la dependencia de las intensidades de STD-NMR con el tiempo de saturación. Con este fin, se procedió a convertir la ecuación 6.2.1 en una serie de Euler.

$$STD_t = STD_{max} \left(- \sum_{j=1}^n \frac{(-k_{sat} t_{sat})^j}{j!} \right) \quad 6.2.3$$

En la ecuación, cuando se extrae el primer término del sumatorio se obtiene una línea recta dependiente del tiempo de saturación con pendiente igual a $dSTD_0$.

$$STD_t = STD_{max} k_{sat} t_{sat} - STD_{max} \left(\sum_{j=2}^n \frac{(-k_{sat} t_{sat})^j}{j!} \right) \quad 6.2.4$$

$$STD_t = dSTD_0 t_{sat} - STD_{max} \left(\sum_{j=2}^n \frac{(-k_{sat} t_{sat})^j}{j!} \right) \quad 6.2.5$$

A partir de la ecuación 6.2.5 se puede afirmar que, a un tiempo de saturación “t” (por ejemplo, $t_{sat} = 2$ s), la saturación de los protones (STD_t) es igual a la velocidad de saturación inicial ($dSTD_0$) más una constante “ K_j^t ”. De manera que, se termina por transformar la ecuación 6.2.1 en la ecuación 6.2.6.

$$STD_t = dSTD_0 t_{sat} + K_j^t \quad 6.2.6$$

La nueva ecuación está compuesta por el término $dSTD_0 \times t_{sat}$, que describe, fundamentalmente, las relajaciones intermoleculares, y por una constante macroscópica (K_j^t), influenciada por los artefactos experimentales. Para explicar los resultados obtenidos en el apartado 5.1.3, se puede suponer que, en la Figura 5.1.2, la combinación de los parámetros químico-físicos utilizados para resolver la matriz CORCEMA, con un $t_{sat}=2$ s, permitieron optimizar tanto el término $dSTD_0 \times t_{sat}$ como el término K_j^t de la ecuación 6.2.6.

En diversos estudios se caracterizó tanto el comportamiento de la saturación en presencia de procesos de reasociación, como su influencia sobre las curvas decrecimiento de STD-NMR (ecuación 6.2.1) (Angulo, Enríquez-Navas et al. 2010, Angulo and Nieto 2011). En sus estudios, Ángulo llegó a relacionar, a partir de la evidencia experimental, que los protones con las señales de STD-NMR más intensas se encuentran más afectados por la reasociación. Además, cuando se intentó obtener la constante de disociación del complejo molecular a partir de las curvas de crecimiento, se encontró que los protones más afectados por la reasociación provocaban una mayor sobrestimación en la determinación de la constante de disociación (Angulo, Enríquez-Navas et al. 2010).

Partiendo de esos resultados se puede afirmar la existencia de una relación entre el proceso de reasociación, la saturación de los protones y la constante de disociación. Por lo tanto, se esperaría que la relación contraria también se cumpliera. Así que, si se utiliza un procedimiento para optimizar la constante de disociación aparente, podríamos ser capaces de predecir de un modo cuantitativo, por medio de la matriz de intercambio conformacional (**K**), los valores de STD-NMR experimentales. De manera que, aunque la aproximación del CORCEMA no tenga en cuenta los artefactos experimentales, y utilicemos valores de STD-NMR experimentales con desviaciones, se esperaría poder rectificar las diferencias en la saturación debidas a las desviaciones mediante la optimización paramétrica de la matriz cinética (**K**) (ecuación 1.2.8) y de relajación (**R**) (ecuación 1.2.9).

La optimización de la constante de disociación permite reproducir aproximadamente tanto los procesos de reasociación como los de relajación T1; y para explicar las relaciones entre ellos se utilizará, como ejemplo, los resultados del trabajo de Angulo. En ese caso particular, la constante de disociación aparente ($K_d^a = 1/K_b^a$) aumenta con el incremento del efecto de la reasociación. Cuando esto ocurre, si se considera que la k_{on} está limitada solamente por la difusión, se espera que también aumente la velocidad de disociación (k_{off}) y además disminuya el tiempo de residencia del ligando en la cavidad de la macromolécula (τ_{res}^u). El menor tiempo de residencia del ligando unido (τ_{res}^u) aumenta la relación $\tau_{res}^l / (\tau_{res}^l + \tau_{res}^u)$, lo que reduce la probabilidad de la aparición de asociaciones múltiples de moléculas de ligando previamente saturadas. Además, este hecho también afectaría la velocidad de la relajación T1. En este caso, se esperaría que el efecto de T1 disminuyese, dado que los mecanismos de relajación del ligando en solución son menos eficientes que los mecanismos del ligando unido a la proteína. Como se comentó en el apartado 5.1.3, el menor tamaño de la red protónica del ligando reduce la velocidad de propagación de la saturación y con ello, también se reducen las velocidades de los mecanismos de relajación T1.

Pero los efectos de la k_{off} no sólo van a modular el intercambio entre estados (matriz **K** del CORCEMA), ya que también, de una manera indirecta, afectan las velocidades de relajación (matriz **R** del CORCEMA). Como se puede constatar en las ecuaciones 1, 12 y 13 del artículo donde fue adaptada la teoría CORCEMA al estudio de la espectroscopia bidimensional de TR-NOESY (Krishna and Moseley 2002), la k_{off} es capaz de modular la autorelajación del ligando (\mathbf{R}_L') y de la proteína (\mathbf{R}_E') cuando están asociados. Como en la mayoría de los complejos moleculares estudiados por STD-NMR, la asociación es débil ($k_{\text{off}} > 1$ Hz), por lo que se puede considerar que la matriz de relajación (**R**) del CORCEMA se comporta como una perturbación de la matriz de intercambio (**K**) (ecuación 5.1.9). Por tanto, con el aumento de $\tau_{\text{res}}^l / (\tau_{\text{res}}^l + \tau_{\text{res}}^u)$, se tendrán menos moléculas de ligando en el sitio, lo que provocaría que el aporte a la saturación de los factores $\mathbf{P}_L \cdot \mathbf{R}_L'$ y $\mathbf{P}_E \cdot \mathbf{R}_E'$ (ecuación 5.1.9) (Jayalakshmi and Krishna 2002) disminuya y, con ello, cambiase la saturación de los protones del ligando. A partir del análisis del CORCEMA, hemos podido apreciar como los procesos de disociación están regulando los términos de relajación del complejo molecular (\mathbf{R}_L' y \mathbf{R}_E'). Pero la K_a y la k_{off} no son los únicos parámetros que permiten ajustar el patrón de STD-NMR. Se conoce, tanto del capítulo Introducción (apartado 1.2.3, ecuación 1.2.6) como del apartado 5.1.3, que los tiempos de correlación modulan la eficiencia y la capacidad de la propagación de la saturación por toda la molécula. De manera que, la variación de los tiempos de correlación influirá directamente en los términos de la matriz de velocidad de relajación completa (**R**) y, por tanto, en los valores de STD-NMR. Por lo tanto, tomando como punto de partida el análisis de la dependencia de los diferentes mecanismos que están regulando los valores de STD-NMR, se puede concluir que, a partir de la optimización de los parámetros químico-físicos (τ_c^p , τ_c^l , K_a y k_{off}) se llegaría a emular las desviaciones en los patrones de STD-NMR.

Para corroborar que el STD-MaPa es capaz de reproducir los patrones de STD-NMR mediante el algoritmo de optimización paramétrica, brevemente se va a comentar la resolución de la estructura de la dimanosa unida al anticuerpo

2G12, a diferentes tiempos de saturación. Con este nuevo ensayo, la metodología implementada en el STD-MaPa demostró su potencial en la resolución de estructuras y en la emulación del patrón de STD-NMR experimental bajo diferentes niveles de interferencias, inducidas tanto por el tiempo de relajación T1 como por la reasociación del ligando.

Para realizar la prueba, se extrajeron todos los datos experimentales que fueron obtenidos por los autores (Enríquez-Navas, Marradi et al. 2011), y las coordenadas de las estructuras de la proteína y de la dimanosa fueron descargados del PDB con el código 1OP3. Para cada tiempo de saturación (0,50, 0,75, 1,00, 1,50, 2,00, 2,50 y 3,00 s) se realizó, con el STD-MaPa, un ciclo de acoplamiento molecular y se optimizaron los tiempos de correlación de la proteína y del ligando, la constante de velocidad de disociación y la constante de asociación. El procedimiento utilizado fue el descrito en el apartados 4.9 y 5.2.6, y se escogieron las estructuras tanto por R_e como por R_t . Los resultados son mostrados en la Tabla 6.2.1, en la Tabla 6.2.2 y en la Figura 6.2.2.

Tabla 6.2.1 Ajuste paramétrico realizado por el STD-MaPa con el factor R_e al resolver la estructura del complejo dimanosa-2G12 en diferentes tiempos de saturación.

Tiempo de saturación	Factor R_e	K_a (M^{-1})	τ_c^p (s)	τ_c^l (s)	k_{off} (s^{-1})
0,50	0,3860	1×10^2	2×10^{-7}	$5,0 \times 10^{-10}$	100
0,75	0,3646	1×10^2	2×10^{-7}	$7,5 \times 10^{-10}$	100
1,00	0,3431	1×10^2	2×10^{-7}	$2,5 \times 10^{-10}$	1×10^3
1,50	0,3051	1×10^2	2×10^{-7}	$2,5 \times 10^{-10}$	1×10^3
2,00	0,2929	1×10^2	2×10^{-7}	$2,5 \times 10^{-10}$	1×10^3
2,50	0,2840	1×10^2	2×10^{-7}	$2,5 \times 10^{-10}$	1×10^3
3,00	0,2721	1×10^2	2×10^{-7}	$2,5 \times 10^{-10}$	1×10^3

Aunque de manera breve, es preciso comentar que el programa STD-MaPa resolvió la estructura descrita por el cristal (PDB: 1OP3) en todos los tiempos de saturación. Tanto en la Tabla 6.2.1 como en la Tabla 6.2.2 se muestran los parámetros optimizados a partir de R_e y R_t . En ellas se observa como el programa modificó, en este caso en concreto, los valores del tiempo de correlación rotacional del ligando (τ_c^l), la constante de asociación (K_a) y la velocidad de disociación (k_{off}), con el aumento del tiempo de saturación (t_{sat}). Como se puede apreciar, los cambios en las constantes llegan a ser de un orden de magnitud. Tanto R_e como R_t redujeron sus valores con el aumento del tiempo de saturación, indicando que, a mayor saturación, las diferencias entre el patrón experimental y el calculado se iban reduciendo. Además, se puede constatar que en todos los casos el factor R_t siempre es menor que R_e ; ya que $W_{cal,k}$ se adaptó mejor que $W_{exp,k}$ a las coordenadas del modelo molecular obtenido con el STD-MaPa.

Tabla 6.2.2 Ajuste paramétrico realizado por el STD-MaPa con el factor R_t al resolver la estructura del complejo dimanosa-2G12 en diferentes tiempos de saturación.

Tiempo de saturación	Factor R_t	K_a (M^{-1})	τ_c^p (s)	τ_c^l (s)	k_{off} (s^{-1})
0,50	0,3440	1×10^2	2×10^{-7}	$2,5 \times 10^{-10}$	1×10^3
0,75	0,3090	1×10^2	2×10^{-7}	$2,5 \times 10^{-10}$	1×10^3
1,00	0,2947	1×10^2	2×10^{-7}	$2,5 \times 10^{-10}$	1×10^3
1,50	0,2820	1×10^2	2×10^{-7}	$2,5 \times 10^{-10}$	1×10^3
2,00	0,2848	1×10^2	2×10^{-7}	$2,5 \times 10^{-10}$	1×10^3
2,50	0,2833	1×10^2	2×10^{-7}	$2,5 \times 10^{-10}$	1×10^3
3,00	0,2674	1×10^3	2×10^{-7}	$2,5 \times 10^{-10}$	1×10^2

El objetivo de este nuevo ensayo era demostrar que el STD-MaPa reconocería la presencia de las distorsiones en los patrones de STD-NMR experimentales. En la Figura 6.2.2 se muestran los patrones de STD-NMR obtenidos con el STD-MaPa para el complejo dimanosa-2G12, en los tiempos de saturación considerados. En la imagen se puede verificar con claridad que el STD-MaPa fue capaz de reproducir los patrones de STD-NMR del complejo dimanosa-2G12 en diferentes tiempos de saturación. Estos resultados muestran además que la respuesta proporcionada por el STD-MaPa se ajusta al incremento de los efectos producidos por los mecanismos de relajación T1 y de reasociación con el aumento del tiempo de saturación. Estos resultados validan la hipótesis de que el STD-MaPa ajusta el patrón de STD-NMR teórico mediante la optimización paramétrica, no afectando el cálculo de las distancias entre el ligando y la proteína. De esta manera, se puede concluir que la resolución estructural mediante el algoritmo que se implementó en el STD-MaPa proporciona resultados cuantitativos y robustos.

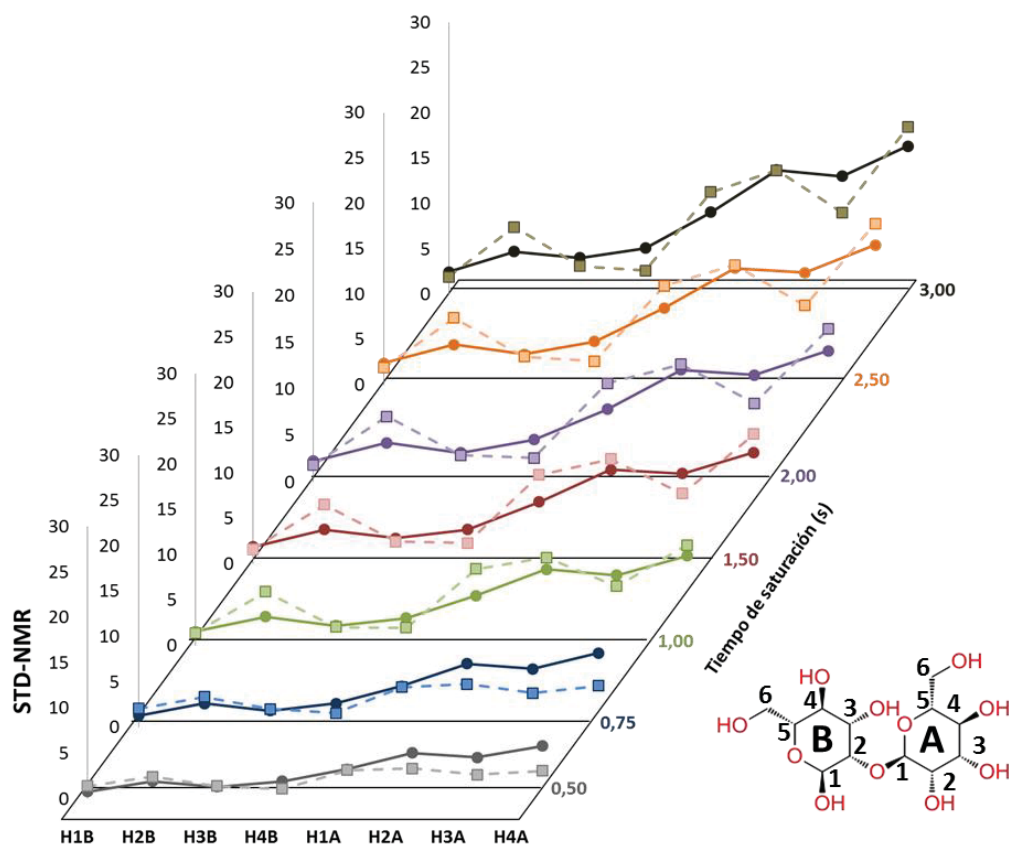


Figura 6.2.2 Patrones de STD-NMR experimentales (líneas sólidas con círculos) y calculados (líneas discontinuas con cuadrados) a partir de la estructura resuelta por el programa STD-MaPa en diferentes tiempos de saturación experimentales. Se representan los patrones de STD-NMR en gris ($t_{\text{sat}}=0,50$ s), azul ($t_{\text{sat}}=0,75$ s), verde ($t_{\text{sat}}=1,00$ s), rojo purpúreo ($t_{\text{sat}}=1,50$ s), violeta ($t_{\text{sat}}=2,00$ s), naranja ($t_{\text{sat}}=2,50$ s) y verde olivo ($t_{\text{sat}}=3,00$ s). En la esquina inferior derecha se encuentra la estructura química de la dimanosa.

Por otro lado, en la resolución mediante el método híbrido de la estructura de la dactilolida unida al tetrámero T_2R (apartado 5.3), se ha supuesto que, con el STD-MaPa, se identificaron cuantitativamente los diferentes modos de unión del compuesto. Sin embargo, aunque es posible la caracterización de múltiples modos de unión a través de experimentos de STD-NMR mediante la aproximación del CORCEMA (Angulo, Díaz et al. 2008, Enríquez-Navas,

Marradi et al. 2011, Muñoz-García, Chabrol et al. 2015), con su aplicación, artefactos como la reasociación de moléculas previamente saturadas pueden impedir la correcta caracterización de los modos de unión (Angulo, Díaz et al. 2008, Angulo and Nieto 2011). Los estudios realizados hasta ahora han utilizado el análisis de pendientes iniciales ($dSTD_0$) para la caracterización cuantitativa de múltiples modos de unión. En ellos, se considera que la pendiente inicial obtenida a partir de los datos experimentales es igual a la suma de las pendientes iniciales de cada modo de unión.

$$dSTD_0^M = \frac{I_0 - I_{sat}^{exp}}{I_0} I_{t_{sat} \rightarrow 0} = \frac{\sum_{i=1}^m (I_0 - I_{sat}^i)}{I_0} I_{t_{sat} \rightarrow 0} = \sum_{i=1}^m dSTD_0^i \quad 6.2.7$$

Sin embargo, la resolución de la estructura de la dactilolida se realizó a partir de los valores de STD-NMR obtenidos en un tiempo de saturación de 2 s (ecuación 5.2.1 del apartado 5.2.7). Aunque se haya validado el STD-MaPa en presencia de efectos provocados por la relajación T1 y la reasociación (Figura 6.2.2), en la resolución estructural de múltiples modos de unión mediante la aproximación matricial del CORCEMA se tiene que cumplir que los modos son independientes entre ellos. Pero los modos son independientes en tiempos de saturación cortos ($t_{sat} \approx 0$) debido a que, en estos tiempos, no se tienen procesos de reasociación. Con tiempos largos, las moléculas que se unen a la macromolécula con una disposición tienen la posibilidad de unirse nuevamente en otra diferente. Cuando esto ocurre, la descripción del STD-NMR a partir de modos de unión independientes deja de ser válida.

En tal caso, las desviaciones del patrón experimental de STD-NMR puede que no se rectifiquen en la optimización paramétrica de cada modo de unión. Para conocer si es válida nuestra aproximación en la resolución del problema de la

reasociación en presencia de múltiples modos de unión, al igual que hicimos con un solo modo de unión, se describieron los valores de STD-NMR con el término de la pendiente inicial ($dSTD_0 \times t_{sat}$) y el término constante (K). Con este fin, se substituyó la ecuación 6.2.6 en la sumatoria de la ecuación 5.2.1.

$$STD_t^M = \sum_{i=1}^m STD_t^i = \sum_{i=1}^m (dSTD_0^i t_{sat} + K_{i,j}^t) \quad 6.2.8$$

La ecuación resultante queda, para un tiempo de saturación “t”, como la sumatoria de la velocidad de crecimiento inicial característica del i-ésimo modo de unión y una constante $K_{i,j}^t$ perturbada por los procesos de reasociación entre los diferentes modos de unión “m”. De manera que se tendrán tantas constantes $K_{i,j}^t$ como modos de unión esté describiendo el ligando. Para separar los efectos de reasociación, la constante $K_{i,j}^t$ se iguala a,

$$K_{ij}^t = K_i^{t'} + \kappa^t \quad 6.2.9$$

y se sustituye en la ecuación 6.2.8,

$$STD_t^M = \sum_{i=1}^m STD_t^i = \sum_{i=1}^m (dSTD_0^i t_{sat} + K_i^{t'}) + \sum_{l=1}^m \kappa_l^t \quad 6.2.10$$

siendo m el número total de modos de unión, $K_i^{t'}$ una constante libre de desviaciones experimentales y los m términos κ_l^t , son constantes que contienen los efectos de reasociación multimodos. Como el término $K_i^{t'}$ no está afectado por los procesos de reasociación, la sumatoria ($\sum_{i=1}^m (dSTD_0^i t_{sat} + K_i^{t'})$) no se

desvía de la aproximación de múltiples modos de unión del CORCEMA. Mientras, κ_i^t incluye todos los efectos experimentales sobre el STD-NMR a tiempo “t”, provocados por las reasociaciones entre diferentes modos de unión. Aunque no se ha comentado anteriormente, se espera que los procesos reasociativos, y con ellos las desviaciones que provocan sobre la saturación de los ligandos, estén regulados por la abundancia entre modos de unión. Al considerar este hecho, se puede afirmar que $(\sum_{l=1}^m \kappa_l^t)$ respeta las relaciones de abundancia entre los diferentes modos de unión. Esto permite reestructurar la ecuación 6.2.10, donde los términos κ_i^t son substituidos por los términos $\kappa_i^{t'}$. Los nuevos términos ($\kappa_i^{t'}$) sólo dependen de los procesos de reasociación entre los mismos modos de unión y por tanto se pueden incluir en el primer sumatorio de la ecuación 6.2.10.

$$STD_t^M = \sum_{i=1}^m STD_t^i = \sum_{i=1}^m (dSTD_0^i t_{sat} + K_i^{t'} + \kappa_i^{t'}) \quad 6.2.11$$

Este nuevo hallazgo, junto con los resultados obtenidos en la Figura 6.2.2 permite afirmar que, mediante la optimización paramétrica de la matriz CORCEMA en cada modo de unión, se pueden describir las reasociaciones multimodos. Por tanto, dado los recientes análisis se puede confirmar que el término κ_i^t (al igual que K_{ij}^t) es emulado por el STD-MaPa mediante la búsqueda de los conjuntos de parámetros (τ_c^p , τ_c^l , K_a y k_{off}) para cada modo de unión. Así que a partir de los cuatro parámetros se van a minimizar los factores R-NOE con la combinación de los patrones de STD-NMR teóricos obtenidos con todos los modos de unión. Como resultado final se tiene que con el software STD-MaPa se llega a rectificar los efectos que se producen en los experimentos de STD-NMR cuando se usan tiempos de saturación largos sin necesidad de obtener una curva de crecimiento experimental.

6.3 STD-MaPa en la resolución de constantes químico-físicas

Posterior al estudio estructural realizado en el apartado 5.2.6, se decidió verificar los ajustes de los parámetros químico-físicos realizados por el STD-MaPa. Para este análisis, fueron tomados en consideración los sistemas formados por la epotilona-A, la podofilotoxina y el compuesto PM060327. A partir de ellos, se analizaron los valores de la constante de asociación obtenidos con el STD-MaPa. Con este fin, primero se obtuvieron experimentalmente las constantes de asociación de estos tres compuestos unidos al heterodímero de α,β -tubulina. Las constantes de asociación de la epotilona-A ($0,8 \pm 0,3 \times 10^4 \text{ M}^{-1}$) y el PM060327 ($K_d = 4,1 \pm 0,9 \times 10^{-8} \text{ M}$; $K_a = 1/K_d = 2 \times 10^7 \text{ M}^{-1}$) se midieron, en nuestro laboratorio, mediante un ensayo de ultracentrífuga analítica por cosedimentación con la tubulina en condiciones no polimerizantes (Pera, Barasoain et al. 2013, Canales, Nieto et al. 2014). En cambio, la constante de asociación de la podofilotoxina ($1,850 \pm 0,007 \times 10^7 \text{ M}^{-1}$) se midió en un ensayo de competición con colchicina por la Dra. Antúnez (Antúnez-Mojica, Rodríguez-Salarichs et al. 2016).

Para analizar la metodología de optimización paramétrica del STD-MaPa, se muestran en la Tabla 6.3.1 los parámetros obtenidos en la resolución estructural de la podofilotoxina- α,β -tubulina, PM060327- α,β -tubulina y epotilona-A- β -tubulina. En esta misma tabla se agrupan todos los parámetros optimizados mediante R_e y R_t para los tres complejos moleculares.

Tabla 6.3.1. Conjunto de parámetros que minimizan los factores R_e y R_t en los complejos cristalográficos de podofilotoxina- α,β -tubulina, PM060327- α,β -tubulina y epotilona-A- β -tubulina.

	Podofilotoxina- α,β - tubulina	PM060327- α,β - tubulina	Epotilona-A- β - tubulina
$\tau_{c\text{-unido}}$ [ns]	40	10^3	40
$\tau_{c\text{-libre}}$ [ns]	0,75	0,25/0,50	0,75
K_a [M^{-1}]	10^8	$10^5/10^4$	$10^4/10^5$
k_{off} [s^{-1}]	10	100	1
R_e/R_t	0,1562/0,1406	0,1740/0,1681	0,2272/0,2055

* Los parámetros obtenidos con los factores R_{tm} , R_{twm} , R_{ms} y R_{msw} no se muestran porque los términos ΔM y ΔM_w , incluidos en los cuatro factores, distorsionan las constantes experimentales.

En estos tres complejos moleculares, la constante de asociación resuelta por el STD-MaPa (tercera fila de la Tabla 6.3.1) mediante los factores R_e y R_t difiere de las constantes obtenidas para los tres complejos mediante ensayos bioquímicos. La metodología experimental utilizada en este trabajo induce una sobreestimación de las constantes de asociación, debido a la presencia de procesos de reasociación y relajación T1. Si se desean eliminar estos efectos se puede utilizar la aproximación propuesta por Mayer y James. Estos autores propusieron registrar curvas de crecimiento de STD-NMR para cada protón y obtener la pendiente inicial de crecimiento mediante un ajuste matemático a una ecuación monoexponencial (Mayer and James 2004). Además, Angulo, basándose en las curvas de crecimiento, propuso un protocolo que eliminaba los artefactos descritos anteriormente y que permitiría determinar correctamente la constante de disociación (Angulo, Enríquez-Navas et al.

2010). Sin embargo, en la resolución estructural de los tres complejos moleculares se utilizó un único tiempo de saturación (ver capítulo de Métodos) y, en estos casos, como se demostró en el apartado 6.2, el STD-MaPa optimiza los valores de K_a para contrarrestar las desviaciones en el patrón de STD-NMR.

Por otro lado, también se analizaron los valores de los tiempos de correlación rotacionales (τ^p_c) de los tres compuestos. En principio, se esperaba que los tres tiempos de correlación fueran iguales debido a que en los tres casos se preparó la misma solución con el heterodímero de α,β -tubulina bajo condiciones despolimerizantes. Sin embargo, los tiempos de correlación (τ^p_c) obtenidos con el STD-MaPa (primera fila de la Tabla 6.3.1) no fueron los mismos para los tres complejos moleculares. En la resolución estructural de los tres complejos moleculares se exploró el rango de τ^p_c de 10 ns a 10^3 ns, en pasos de 10 ns. Como se puede observar en la Tabla 6.3.1, se obtuvo un tiempo de correlación para la proteína de 40ns en el sistema con podofilotoxina y con epotilona-A. Sin embargo, el τ^p_c para el sistema con PM060327 fue 10^3 ns. Aunque se esperaba que los τ^p_c entre los tres complejos tuvieran ciertas diferencias debido al sobreajuste que realiza el STD-MaPa para emular las anomalías en el patrón de STD-NMR, para el mismo tiempo de saturación y con sistemas tan similares, no se esperaban cambios de dos órdenes de diferencia. Así que los cambios en los tiempos de correlación del heterodímero de α,β -tubulina son debidos a otros factores.

En un estudio previo, ya se había calculado el tiempo de correlación del dímero de tubulina a través de la aproximación brindada por el HYDROPRO (Canales, Rodríguez-Salarichs et al. 2011). La aproximación del HYDRPRO considera que las moléculas son globulares y con este método se determinó que el τ^p_c del heterodímero de α,β -tubulina era 60 ns. Pero la tubulina es un dímero formado por dos proteínas globulares, cada una de ~50 kDa, que en su conjunto no forman un compuesto globular. Así que la desviación que tiene la α,β -tubulina del modelo globular considerado por el HYDROPRO puede inducirle a éste a

calcular un radio hidrodinámico mayor que el real. Si se tiene en cuenta esta desviación por parte del HYDROPRO, se puede decir que su resultado ($\tau^p_c=60$ ns) se asemeja al resultado obtenido con el STD-MaPa para los complejos con podofilotoxina ($\tau^p_c=40$ ns) y epotilona-A ($\tau^p_c=40$ ns) (Tabla 6.3.1).

En otro estudio, en el que se analizaba el equilibrio de plegamiento de los heterodímeros de α,β -tubulina en presencia de cloruro de guanidinio (GndCl), mediante técnicas de Dicroísmo Circular (CD), Espectroscopia de Correlación de Fluorescencia de dos fotones (FCS) y Espectroscopia de Fluorescencia Resuelta en el Tiempo, se llegó a determinar por fluorescencia el tiempo de correlación rotacional de la tubulina (Sánchez, Brunet et al. 2004). Para dicho estudio, la tubulina utilizada se extrajo de cerebro bovino (Andreu, Oliva et al. 2002) siguiendo el mismo protocolo usado en la extracción de tubulina en los experimentos de RMN con los complejos de podofilotoxina- α,β -tubulina, PM060327- α,β -tubulina y epotilona-A- α,β -tubulina (aparatado 4.3). Además, la viscosidad del tampón utilizado en este estudio es similar a la viscosidad de las soluciones preparadas para los experimentos de RMN con la podofilotoxina, el PM060327 y la epotilona-A; de manera que los resultados obtenidos en este estudio podrían validar los resultados del STD-MaPa. Finalmente, en este estudio el tiempo de correlación de la proteína fue medida por fluorescencia y con diferentes concentraciones de GndCl (0-0,5 M). Los resultados mostraron que en un rango de concentración de GndCl de 0 a 0,5 ns, el tiempo de correlación del heterodímero de α,β -tubulina está comprendido entre los 40 ns ($c(\text{GndCl})=0$ M) y los 48 ns ($c(\text{GndCl})=0,5$ M). Así que, en ausencia de GndCl, los dos métodos obtienen el mismo valor de τ^p_c para el heterodímero de α,β -tubulina. Este resultado demuestra que el STD-MaPa, mediante los factores R_e y R_t , optimizó correctamente el tiempo de correlación del dímero de α,β -tubulina en presencia de podofilotoxina y epotilona-A. Además, indicaría que los cálculos realizados por el HYDROPRO se desviaron 20 ns; desviación que probablemente fue producida por la aproximación geométrica que utiliza el programa.

Hasta el momento, los resultados con el STD-MaPa para podofilotoxina y epotilona-A se han podido validar. Sin embargo, no se puede obviar que, con el mismo protocolo, el programa obtiene un tiempo de correlación rotacional, para el dímero de α,β -tubulina, de $1\mu\text{s}$ en el complejo de PM060327- α,β -tubulina. Para conocer los motivos que produjeron esta gran desviación, se empezó por calcular la masa molecular correspondiente a un tiempo de correlación rotacional de $1\mu\text{s}$. Con este fin y mediante una aproximación grosera, se determinó la masa molecular de una molécula con un τ_c de $1\mu\text{s}$, a 25°C y en una solución de deuterio (Cho, Urquidi et al. 1999, Cavanagh, Fairbrother et al. 2010). El cálculo de τ_c se realizó mediante una aproximación empírica basada en la Ley de Stokes y, según sus resultados, la masa molecular de una molécula con un τ_c^p de $1\mu\text{s}$ es 1666,944 kDa. Esto equivale a decir que la tubulina está formando aglomerado u oligómero compuestos por ~ 15 dímeros.

En el trabajo (Pera, Barasoain et al. 2013) sobre PM060327, se analizó mediante ultracentrífuga analítica, el estado de polimerización presente en las condiciones de los experimentos de RMN. Los resultados mostraron que solo el 93% de la tubulina estaba formando dímeros, mientras que el resto se encontraba en diversos estados de polimerización. Además, en el mismo trabajo, se llegó a documentar la presencia de anillos dobles con un diámetro medio de 38 nm; y según estudios previos, están compuestos entre 12 y 16 heterodímeros de α,β -tubulina (Díaz, Pantos et al. 1994).

Según las evidencias mostradas hasta este momento, el programa STD-MaPa detectó el grado de polimerización de la tubulina y, en consecuencia, corrigió el valor del tiempo de correlación rotacional τ_c^p pasando éste de 40 ns a $1\mu\text{s}$. Por lo tanto, el algoritmo de optimización paramétrica del STD-MaPa tiene la capacidad de reconocer el tamaño de los compuestos y calcular correctamente el tiempo de correlación rotacional. Además, a partir de los análisis realizados en este apartado y los ajustes paramétricos obtenidos con el STD-MaPa, en la resolución del complejo de dimanosa-2G12 a diferentes tiempos de saturación (en el apartado 6.2, la Tabla 6.2.1 y Tabla 6.2.2), se concluye que, en

presencia de desviaciones en el patrón de STD-NMR experimental, los parámetros de la matriz de velocidad de intercambio conformacional completo (**K**) (apartado 1.2.4, ecuación 1.2.9) se verán más afectados que los parámetros de la matriz de velocidad de relajación completa (**R**) (apartado 1.2.4, ecuación 1.2.8). Finalmente, a partir de las evidencias expuestas se puede concluir que con la metodología implementada en el programa STD-MaPa se puede llegar a optimizar los tiempos de correlación de las proteínas, así como sus estados de polimerización.

6.4 Descripción del mecanismo de unión covalente de la dactilolida y sus derivados a la β -tubulina

Uno de los objetivos de este trabajo era la obtención de un método híbrido que permitiera, mediante la combinación de diversas técnicas experimentales, obtener las coordenadas reales de un complejo proteína-ligando. De esta manera, en el apartado 5.3 se resolvió la estructura de la dactilolida unida al complejo T₂R mediante el método híbrido de RMN y DRX. Los resultados indicaron que la molécula se asoció con al menos dos orientaciones a la β -tubulina. A partir de la caracterización estructural del complejo molecular, se identificó uno de los modos de unión como el más abundante (83% de la dactilolida enlazada) y, según su disposición en el sitio, ésta sería la estructura inicial de la reacción covalente con la β -tubulina.

Si se superponen las coordenadas de β 1- y β 2- tubulina unidas a la dactilolida (modo predominante del método híbrido) y la zampanolida (PDB: 4I4T) respectivamente, se observa como ambos compuestos presentan la misma orientación en la cavidad de β -tubulina (Figura 6.4.1). En la imagen se puede ver como la zampanolida interacciona a través de la cadena lateral con el lazo-M, siendo éste el mecanismo que mejora la estabilidad de la unión de la zampanolida con respecto a la dactilolida. Mientras, la dactilolida se entierra

más en el sitio en ausencia de esta fuente de estabilización (Figura 6.4.1-A) y se estabiliza mediante la oclusión del anillo 4-metilen-tetrahidropirano entre la hélice H7 y el lazo B9-B10. Estas circunstancias favorecen los procesos de disociación rápidos ($k_{\text{off}}=10 \text{ s}^{-1}$), lo que induciría una disminución en la velocidad de la reacción covalente con la β -tubulina.

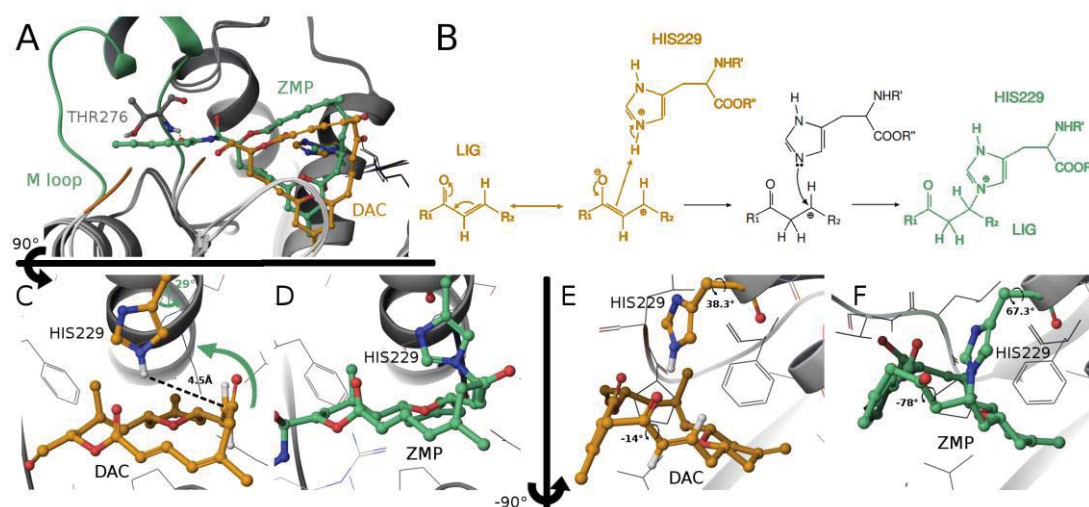


Figura 6.4.1. Representación del mecanismo de reacción covalente de la dactilolida y la zampanolida en el sitio interno de la β -tubulina. El mecanismo de reacción covalente está representado en tres pasos que se diferencian con el naranja (estructuras antes de la reacción covalente), negro (estructura intermedia) y verde (estructura posterior a la reacción covalente). (A) Superposición de la estructura de $\beta 1$ - y $\beta 2$ - tubulina en presencia de DAC (naranja) y ZMP (verde) respectivamente. (B) Representación, según la Teoría de Enlace de Valencia (TEV), del mecanismo de la reacción covalente entre el macrociclo de ZMP o DAC con el residuo de histidina 229. En naranja, se representan los pasos que están relacionados con la estructura del complejo de DAC- T_2R obtenida mediante el método híbrido. En verde, se representan los pasos que están relacionados con la estructura del complejo de ZMP- T_2R obtenida mediante DRX. En negro, el paso intermedio del mecanismo de unión

covalente. (C y D) Cambio de perspectiva de la imagen “A” mediante un giro de 90° con respecto un eje horizontal. (E y F) Cambio de perspectiva de las imagenes “C” y “D” respectivamente mediante un giro de -90° con respecto a un eje vertical.

Sin embargo, debido a la similitud estructural entre los dos compuestos y a la disposición espacial adquirida en la unión a la β -tubulina, se esperaría que el mecanismo de la reacción covalente de ambas fuese el mismo. Así que, teniendo en cuenta esta consideración, se pudo caracterizar el mecanismo de reacción covalente de esta familia de compuestos en la unión con el sitio interno de β -tubulina, a partir del análisis de las estructuras de los complejos moleculares con la dactilolida y la zampanolida.

Una de las características más conspicua del sitio interno de la β -tubulina es la elevada exposición que presenta la cadena lateral del residuo HIS229 (Lowe, Li et al. 2001). Este residuo, en presencia de un compuesto (Lowe, Li et al. 2001, Matesanz, Trigili et al. 2015), se suele encontrar en el medio de la hélice H7, dividiendo la cavidad a la mitad. Además, según diversos estudios sobre la protonación del anillo de imidazol de la HIS229, se ha llegado a identificar que éste se encuentra doblemente protonado ($pK_{HIS229}=7,2$ a $pH=6.5$) (Matesanz, Barasoain et al. 2008, Matesanz, Trigili et al. 2015).

Se conoce que tanto la dactilolida como la zampanolida reaccionan covalentemente con la HIS229 a través del éster α,β -insaturado del macrociclo, mediante una adición de Michael (Figura 6.4.1-B). Inicialmente, la reacción covalente de los dos compuestos con la HIS229 fue caracterizada mediante estudios de Espectroscopia de Masas y Modelado Molecular (Field, Pera et al. 2012); y un año más tarde, esta información se confirmó con la resolución estructural de la zampanolida y el complejo T₂R mediante técnicas cristalográficas (Prota, Bargsten et al. 2013). De manera que se esperaría que el estado de protonación de la HIS229 influyese en el mecanismo de reacción covalente.

En la HIS229, debido al estado de protonación del residuo y las características electronegativas del nitrógeno, se crea un efecto inductivo que desplaza la densidad electrónica del enlace σ localizado entre NE2 y HE2 hacia el nitrógeno NE2. Este efecto electrónico aumenta la reactividad de la HIS229, cargando positivamente el protón HE2. Dada la importancia de esta distribución de carga, recientemente fue calculada la distribución electrónica y el potencial electrostático de la HIS229 mediante una aproximación mecano-cuántica (Tang, Rodríguez-Salarichs et al. 2017). A partir de estos cálculos, se encontrarían claras evidencias de la localización electrónica sobre los nitrógenos del anillo de imidazol, y de la presencia de un potencial electrostático positivo sobre HE2 (Figura 3-E, de (Tang, Rodríguez-Salarichs et al. 2017)).

Si se observa la Figura 6.4.1-C, se puede apreciar que, en la unión no covalente de la DAC, el residuo de histidina 229 (HIS229) se orienta hacia el centro del macrociclo, maximizando su contacto con toda la estructura del compuesto. Cuando la HIS229 adquiere esta orientación con respecto a la DAC, la distancia entre el hidrógeno HE2 del residuo HIS229 y el grupo etileno del macrociclo es sólo de 4.5 Å. Con esta distancia, la densidad electrónica localizada sobre el enlace π conjugado del macrociclo (ZMP o DAC) favorece el acercamiento (Figura 6.4.1-B, paso de la reacción representado en naranja) y posterior transferencia del hidrógeno HE2, desde el residuo 229 hacia el ligando (Figura 6.4.1-B, paso de la reacción representado en negro).

Debido a la atracción electrostática existente entre el grupo etileno (DAC o ZMP) y el átomo HE2 del residuo, el anillo de imidazol (HIS229) se orienta, mediante un giro de 29°, hacia el éster α,β -insaturado del compuesto (DAC o ZMP) (Figura 6.4.1-C, E y F). Además, el macrociclo (DAC y ZMP) tuerce su estructura 64°, induciendo un desacoplamiento en la conjugación entre el grupo éster y el etileno (Figura 6.4.1-C, D y F). Estos cambios conformacionales provocan un acercamiento entre el residuo (HIS229) y el compuesto (DAC o ZMP), favoreciendo la ruptura del enlace σ entre NE2 y HE2 y la transferencia del átomo de HE2 hacia el macrociclo Figura 6.4.1-B (estructura representada

en negro). Este suceso transfiere la carga positiva de la HIS229 hacia el macrociclo e induce un posterior ataque nucleofílico del átomo NE2 (HIS229) sobre el carbono β del compuesto (carbono que tiene la carga positiva en la Figura 6.4.1-B). Finalmente, con posterioridad a todos los cambios torsionales y transferencias, se obtiene el compuesto (DAC y ZMP) unidos covalentemente a la HIS229 (Figura 6.4.1-B (estructura representada en verde), D y F).

Con el análisis del mecanismo de unión covalente de la DAC y la ZMP se concluye que éste depende, principalmente, de la cercanía entre la HIS229 y el grupo insaturado del ligando. De manera que la orientación del ligando en la cavidad de β -tubulina debe permitir que el grupo químico, que vaya a ser partícipe de la adición conjugada con la HIS229, se localice entorno al átomo HE2 del residuo. Otro factor que favorece la reacción covalente es la estabilidad de la asociación previa a la reacción. Se llegó a esta conclusión porque, en la caracterización de la unión no covalente de la zampanolida y la dactilolida por técnicas bioquímicas realizada en el apartado 1.3 de la Introducción, se determinó que la constante de asociación aparente de la zampanolida es $K_a^{ap}(26\text{ }^{\circ}\text{C})=1,4 \pm 0,3 \times 10^8\text{ M}^{-1}$, mientras que en la dactilolida se obtuvo un valor de $K_a^{ap}(26\text{ }^{\circ}\text{C})=0,6 \pm 0,1 \times 10^6\text{ M}^{-1}$. También se explicó en ese apartado que la diferencia, de dos órdenes de magnitud entre las constantes, se debía a la presencia de la cadena lateral en la ZMP; ya que ésta estabilizaba la unión de la molécula mediante su interacción directa con el lazo-M de β -tubulina. Si se comparan los tiempos de las reacciones covalentes entre la zampanolida ($t \approx 30\text{ min}$) y la dactilolida ($t \approx 8\text{ hr}$), se puede comprobar cómo la unión covalente ocurre de manera más rápida en la zampanolida. Es decir, la estabilización de los compuestos favorecerá la velocidad de la reacción covalente. Este hecho se debe a que la reacción covalente compite con la disociación del ligando (k_{off}); y si el complejo molecular es poco estable, se favorecerá su disociación, como ocurre con la dactilolida.

Como resultado del análisis se llegó a explicar la estabilización de la asociación no covalente como el paso limitante en la reacción de la dactilolida y la β -

tubulina. Debido a que, en este paso, la interacción de la cadena lateral con el lazo-M favoreció la estabilización de los compuestos (Matesanz, Rodríguez-Salarichs et al. 2011, Prota, Bargsten et al. 2013, Matesanz, Trigili et al. 2015) y por tanto, su reacción covalente. Por este motivo, el análisis realizado sobre el mecanismo de reacción de la ZMP y la DAC aportó una valiosa información que bien puede ser aplicada en el diseño de nuevos compuestos covalentes dirigidos contra el sitio del Taxol®.

6.5 Resolución de la estructura de complejos moleculares con estructuras similares por STD-MaPa

Aunque los métodos desarrollados en este trabajo se han orientado hacia la determinación analítica y cuantitativa de las estructuras de los complejos moleculares, la metodología implementada en el STD-MaPa también puede utilizarse en la resolución aproximada de las estructuras de los complejos moleculares en ausencia de la información estructural de la macromolécula. Con este fin, se pueden tomar como punto de partida las coordenadas de la macromolécula unida a otro compuesto, el cual tendría una alta similitud estructural con el compuesto en estudio. De esta manera, con el STD-MaPa y diversos algoritmos de exploración conformacional y optimización estructural, se pudo llegar a resolver la estructura descrita por el patrón de STD-MaPa. Pero esta aproximación tiene el inconveniente de no controlar, por falta de información experimental, si las coordenadas obtenidas son las que mejor describirían el patrón de STD-MaPa. Sin embargo, si la estructura del compuesto utilizada como plantilla tiene una alta similitud, remodelar la estructura del receptor y adaptarla a la estructura del complejo en estudio pudiese constituir una aproximación válida.

Con el fin de verificar esta aproximación, en este apartado se va a comentar la resolución, con el STD-MaPa, de las estructuras de los complejos moleculares

formados entre el dímero de α,β -tubulina y, por una parte, la azatilon-2, que es un derivado sintético de la familia de azatilonas y presenta una alta similitud estructural con la epotilon-A (Jantsch, Nieto et al. 2016) (Figura 6.5.2); y, por otra, dos análogos naturales de la podofilotoxina extraídos de la planta medicinal Mexicana *Bursera fagaroides* (Antúñez-Mojica, Rodríguez-Salarichs et al. 2016) (Figura 6.5.1). Para realizar estos trabajos, se utilizaron las coordenadas de las proteínas extraídas de las estructuras cristalográficas de la epotilon-A (PDB: 4I50) y la podofilotoxina (PDB: 1SA1), respectivamente. Con estas estructuras, mediante el STD-MaPa y diversos algoritmos de exploración conformacional, en ambos trabajos se llegaron a describir los datos de STD-NMR experimentales de los complejos de azatilon-2- β -tubulina (Figura 6.5.2) y dos análogos naturales de la podofilotoxina unidas al heterodímero de α,β -tubulina (Figura 6.5.1).

Entre las dos imágenes se muestran las estructuras químicas de la epotilon-A, la azatilon-2, la podofilotoxina y sus derivados 2 y 3. En ellas, se puede ver que la podofilotoxina y el derivado 2 son los compuestos que presentan mayor similitud estructural; ya que sus estructuras solo se diferencian en la valencia de su enlace 7'-8'. Sin embargo, la epotilon-A y la azatilon-2 presentan diferencias en la estructura que pueden afectar la correcta caracterización de la unión a β -tubulina (Figura 6.5.2, grupos químicos representados en rojo y azul). Las diferencias se ubican en regiones flexibles de la molécula que, según la estructura cristalográfica de la epotilon-A (Prota, Bargsten et al. 2013), son importantes en la unión al sitio del Taxol®.

En el protocolo que se siguió para ambos casos, se llevó a cabo un primer paso en el que se generaron las primeras orientaciones de la azatilon-2 (Figura 6.5.2-A) y los derivados 2 y 3 con el STD-MaPa; usando como plantillas las estructuras cristalográficas de la epotilon-A y la podofilotoxina, respectivamente. En este paso, los modos de unión escogidos por el algoritmo para la azatilon-2 y los derivados 2 y 3, fueron similares a los descritos por las estructuras cristalográficas de la epotilon-A y la podofilotoxina. Este resultado

era de esperar, ya que en ambos casos la similitud estructural es muy alta y las superficies de los dos sitios de unión estaban predisuestas al reconocimiento específico de la epotilona-A y la podofilotoxina.

En un segundo paso, debido a que las estructuras de los derivados y la podofilotoxinas son rígidas y, además, las diferencias entre las estructuras son muy pequeñas, se consideró apropiada la optimización estructural mediante el Macromodel (Mohamadi, Richards et al. 1990). Los resultados fueron superpuestos con la estructura de la podofilotoxina y representados en la Figura 6.5.1. En la imagen se observa cómo se logra adaptar la estructura del sitio de unión de la podofilotoxina para la asociación de los derivados 2 y 3 (Figura 6.5.1-B y C). Además, se puede constatar que los patrones de STD-NMR obtenidos se asemejan a los experimentales con un $R_e=0,37$ para el compuesto 2 y un $R_e=0,23$ para el compuesto 3.

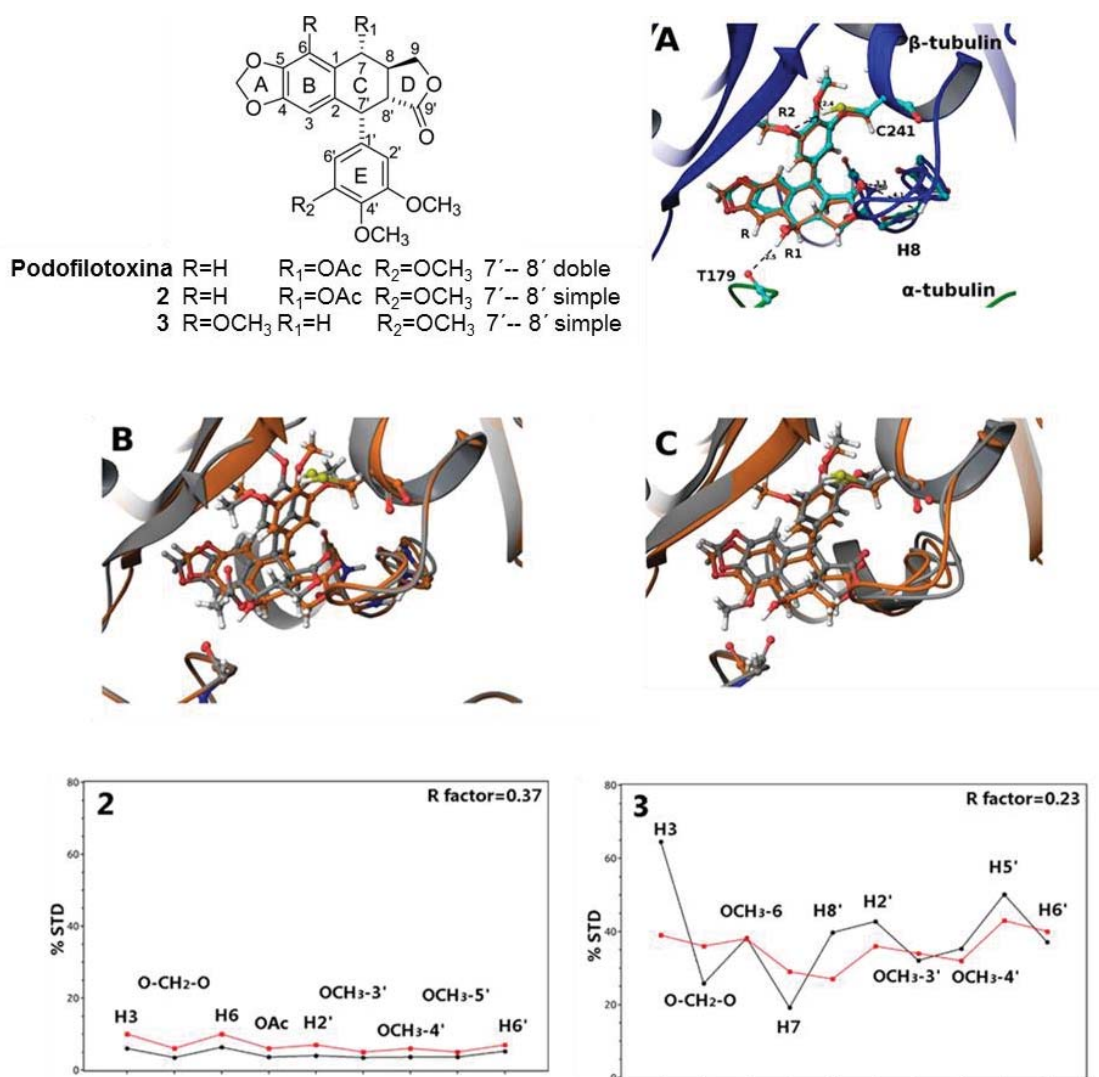
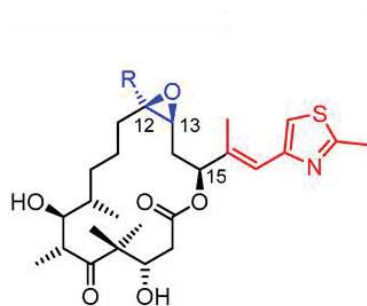


Figura 6.5.1 . Representación de los modos de unión de la podofilotoxina y sus dos análogos 2 y 3. La podofilotoxina se representa en naranja y sus derivados en gris. (A) Superposición de la estructura cristalográfica de la podofilotoxina PDB 1SA1 (azul claro) y la estructura obtenida con el STD-MaPa (naranja). (B) Superposición de la estructura de la podofilotoxina y el compuesto 2. (C) Superposición de la estructura de la podofilotoxina y el compuesto 3 (conformación A de la Figura 8 de (Antúñez-Mojica, Rodríguez-Salarichs et al. 2016)). En el inferior de la imagen, se encuentra la comparación entre los patrones de STD-NMR experimentales y los calculados a partir de los modelos moleculares obtenidos con el STD-MaPa de los compuestos 2 y 3.

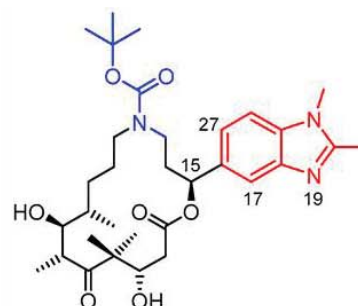
Sin embargo, los valores del factor R_e son, tanto en el derivado 2 como en el 3, ligeramente elevados. Además, el patrón de STD-NMR teórico del compuesto 3, aunque presenta un R_e menor, describe accidentalmente el patrón de STD-NMR experimental. Con todos estos resultados, aunque se llegó a describir correctamente la orientación de los compuestos según sus patrones de STD-NMR y los valores de R_e , se debería someter a las dos estructuras a un nuevo paso de refinamiento estructural guiados por RMN. De esta manera, se podría minimizar el valor del factor R_e y reducir las diferencias entre los patrones de STD-NMR de los dos modelos moleculares (compuesto 2 y 3) y sus patrones experimentales.

Por otro lado, como las regiones que diferencian a la azatilon-2 son flexibles, se optó por realizar una exploración conformacional intensiva (Figura 6.5.2). Con este estudio, se obtuvieron 144 conformaciones de azatilon-2, mediante la rotación de los enlaces C15-C16 y O-C(CH₃)₃ y una posterior optimización local para reducir los choques estéricos. Todas las estructuras resultantes se evaluaron con el factor R_e y sus valores se representaron en la Figura 6.5.2-B.

Según el mapa obtenido con el factor R_e , la azatilon-2 tiene dos conformaciones que minimizan por igual el factor R_e (Figura 6.5.2-B) y además, éstos exhiben el mismo patrón de STD-NMR (Figura 6.5.2-C y D). Por otro lado, a partir del análisis del mapa se pudo afirmar que la rotación del terc-butilo (estructura química representada en azul en la Figura 6.5.2) perturba poco el patrón de STD-NMR (Figura 6.5.2-B). Mientras, la cadena lateral de la azatilon-2 (estructura química representada en rojo en la Figura 6.5.2), tiene dos orientaciones posibles, las cuales son coplanares a la cadena lateral de la epotilon-A en la estructura cristalográfica (Figura 6.5.2-G y H).



Epotilona-A: R = H



Azatilon 2

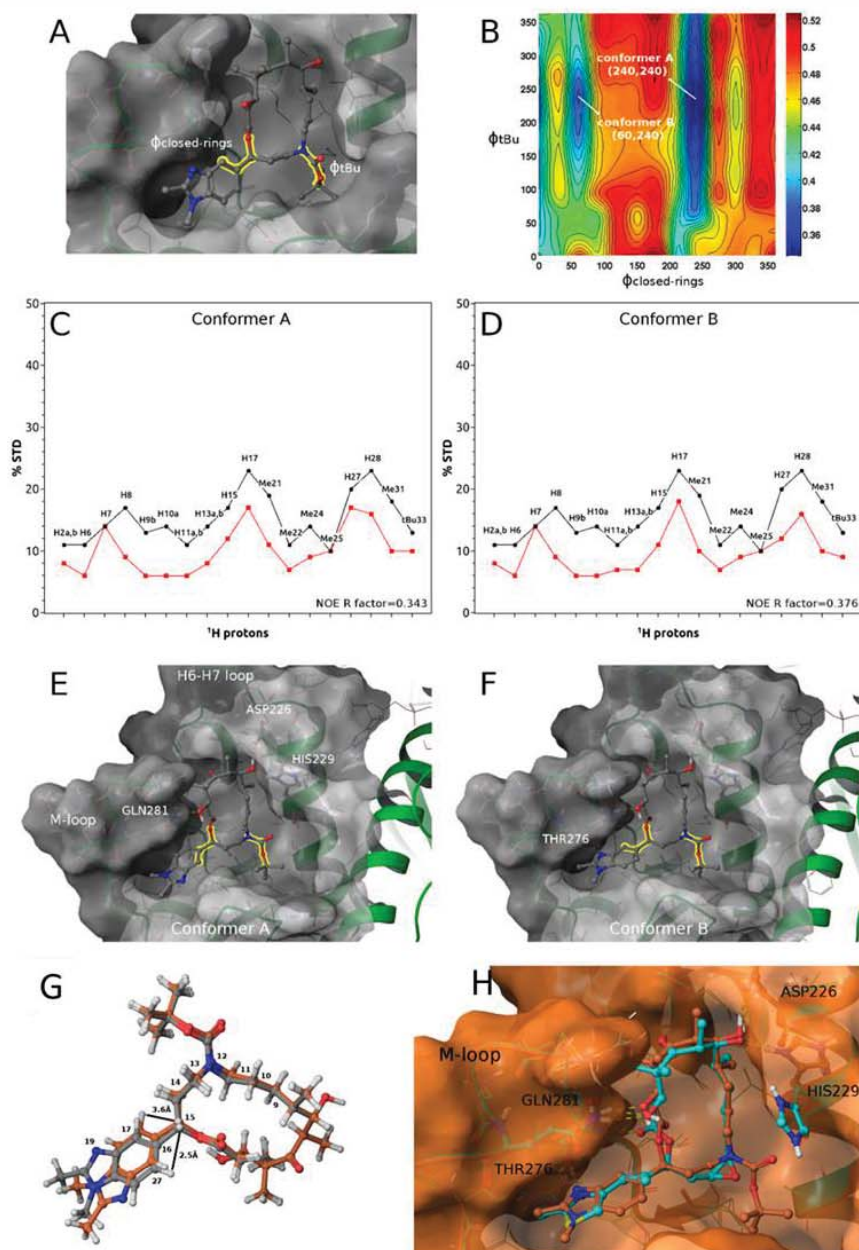


Figura 6.5.2 Representación de las estructuras químicas de la epotilona-A y la azatilon-2. Las diferencias que presentan en sus estructuras químicas están representadas en rojo y azul. (A) Representación de la azatilon-2 unida al sitio interno de la β -tubulina. Las coordenadas fueron obtenidas con un primer ciclo con el STD-MaPa. (B) Superficie del factor R-NOE obtenida a partir de la exploración conformacional con el Macromodel y el STD-MaPa. Mediante la rotación de los enlaces C15-C16 y O-C(CH₃)₃ se obtuvieron 144 conformaciones de la azatilon-2. Los enlaces que se rotaron están resaltados en amarillo en la imagen "A". (C) Patrón de STD-NMR experimental (negro) y teórico (rojo) del conformero A. (D) Patrón de STD-NMR experimental (negro) y teórico (rojo) del conformero B. (E) Modelo refinado del conformero A unido a la β -tubulina. (F) Modelo refinado del conformero B unido a la β -tubulina. (G) Superposición de las estructuras del conformero A (naranja) y B (gris). (H) Superposición del modelo 3D del complejo azatilon-2- β -tubulina (conformero B en naranja) y el complejo epotilona-A- β -tubulina (azul claro).

Aunque el estudio permitió caracterizar la unión de la azatilon-2 al igual que con la podofilotoxina y sus derivados, los patrones de STD-NMR teóricos no se encuentran bien descritos. Si comparamos los valores de R_e de la podofilotoxina ($R_e=0,141$) y la epotilona-A ($R_e=0,206$) obtenidos en el apartado 5.2.6 con el STD-MaPa con los calculados para el compuesto 2 ($R_e=0,37$), el 3 ($R_e=0,23$) y la azatilon-2 ($R_e=0,343$), se concluye que es necesario un nuevo refinamiento para las estructuras de los tres complejos.

A partir del breve análisis realizado en estos sistemas moleculares, se puede concluir que el STD-MaPa es también útil en la descripción de derivados estructuralmente cercanos a compuestos previamente resueltos mediante otras técnicas experimentales. En estos casos, el modelo molecular obtenido será una aproximación grosera a la estructura real; pero permitirá la caracterización estructural de los complejos moleculares cuando no se tenga información sobre la estructura del receptor.

Conclusiones

7 Conclusiones

En el trabajo, se han diseñado cinco alternativas al factor R_e (R_t , R_{tm} , R_{twm} , R_{ms} y R_{msw}) dirigidas a resolver estructuras de complejos moleculares a partir de datos de Diferencia de Transferida de Saturación. En el diseño de las alternativas se combinaron los términos $W_{cal,k}$, ΔM , ΔM_w , $N_i \Delta M^2$ y $N_i \Delta M_w^2$ para mejorar la adaptación del factor R-NOE al proceso de modelado molecular y para reducir su dependencia paramétrica. Además, se obtuvo la ecuación de error de los factores R_e y R_t para evaluar la significación estadística de los resultados. A partir de los análisis realizados con los seis factores R-NOE (R_e , R_t , R_{tm} , R_{twm} , R_{ms} y R_{msw}), los factores se separaron en dos grupos según su utilidad. El primer grupo está constituido por R_{tm} , R_{twm} , R_{ms} y R_{msw} , es poco sensible a los cambio paramétricos y localiza la estructura real entre cinco candidatos, a partir de un conjunto de modelos moleculares e independientemente de los parámetros químico-físicos utilizados. El segundo grupo está formado por R_e y R_t , los dos factores mostraron ser idóneos en los procesos de optimización estructural y paramétrica. Aunque los dos factores R-NOE tuvieron un rendimiento similar, R_t se adaptó mejor al proceso de modelado molecular, al reducir más su valor con respecto a R_e , en presencia de los modelos reales o modelos estructuralmente cercanos al real.

Además, se diseñó una metodología automática de acoplamiento molecular basada en datos de Diferencia de Transferencia de Saturación. La metodología fue implementada en el programa STD-MaPa y permite resolver las estructuras de los complejos moleculares utilizando los seis factores R-NOE de manera activa en la evaluación de los modelos moleculares generados por el propio programa. El STD-MaPa se puede ejecutar a través de una interfaz visual amigable que permite su utilización para personal no experto.

Las diferentes partes de la metodología implementada en el programa STD-MaPa fueron probadas. Los resultados mostraron que el programa el 100% de las veces genera la estructura que describen los datos de STD-NMR y la

identifica utilizando los factores R-NOE como funciones de evaluación. Estos resultados se obtuvieron bajo ciertas condiciones que aseguran la fiabilidad de los resultados.

El STD-MaPa, a partir de los datos de STD-NMR y las coordenadas del complejo molecular calcula mediante un algoritmo de ajuste paramétrico la constante aparente de asociación, la constante de velocidad aparente de disociación, el tiempo de correlación aparente del ligando y mediante un archivo de órdenes (script) acoplado al STD-MaPa el tiempo de correlación aparente de la macromolécula. Así se obtuvo el tiempo de correlación del heterodímero de α,β -tubulina y se identificó su estado polimérico al caracterizar su unión con el PM060327.

Dada la flexibilidad de la metodología automática implementada en el STD-MaPa, permite su adaptación a métodos externos para la exploración espacial con el uso activo de patrones de Diferencia de Transferencia de Saturación experimentales. En nuestro trabajo, mediante el procesamiento automático implementado en el STD-MaPa, se caracterizó el mapa de los seis factores R-NOE en el estudio conformacional de los complejos de kifunesina-GMII, salacinol-GMII y epotilona-A- α,β -tubulina. A partir de los resultados de este estudio se identificaron dos conformaciones en la epotilona-A que coexisten en la unión al heterodímero de α,β -tubulina.

Por otra parte, para aplicar el STD-MaPa y el CORCEMA-ST en el cálculo de múltiples modos de unión, las Ecuaciones Generales de Concentraciones para el CORCEMA (EGC-CORCEMA) fueron resueltas. Las ecuaciones obtienen los datos de concentraciones adaptadas a los cálculos con la aproximación cuantitativa CORCEMA del acoplamiento molecular de múltiples orientaciones y relaciones de proporcionalidad. Además, se comprobó que con el STD-MaPa no se necesita obtener una curva de crecimiento en función del tiempo de saturación para eliminar ciertos artefactos experimentales. En el estudio de la unión de la dimanosidasa y la proteína 2G12, los patrones de Diferencia de Transferencia de Saturación experimentales para tiempos de saturación de

0,50, 0,75, 1,00, 1,50, 2,00, 2,50 y 3,00 segundos se obtuvieron con el STD-MaPa y las estructuras cristalográficas. Ya que, el STD-MaPa rectifica las desviaciones debidas a la reasociación y a los procesos de relajación de la magnetización (T_1) con los factores R_{tm} , R_{twm} , R_{ms} y R_{msw} y el sobreajuste de la constante de asociación y la velocidad de disociación, permitiendo la caracterización cuantitativa de las coordenadas 3D de los complejos moleculares

El STD-MaPa permitió identificar la orientación de derivados de la familia de epotilonas y podofilotoxinas en el sitio de unión del complejo molecular utilizando como plantilla las coordenadas de la proteína diana de compuestos con alta similaridad estructural. Estos resultados se alcanzaron mediante el acoplamiento del STD-MaPa con diversas metodologías externas de exploración conformacional y optimización estructural que permitieron, mediante la reorientación de los residuos de las proteínas, alcanzar a describir un sitio de unión adaptado al acoplamiento de los compuestos estudiados. Aunque no contamos con herramientas para asegurar que las estructuras obtenidas sean las que mejor se adaptan a los datos de Diferencia de Transferencia de Saturación, el algoritmo utilizado permitió obtener modelos moleculares con patrones de STD-NMR semejantes a los datos experimentales.

Posterior a todas las pruebas que se realizaron sobre STD-MaPa ratifican que es capaz de resolver el acoplamiento de un complejo molecular descrito por los datos de STD-NMR, diseñamos una metodología híbrido que permite resolver las coordenadas 3D de un complejo molecular proteína-ligando mediante el acoplamiento de las técnicas de TR-NOESY, STD-NMR y DRX. Con el nuevo método híbrido se resolvió la estructura del tetrámero T_2R describiendo una unión reversiblemente con la dactilolida. Los resultados mostraron que la dactilolida se une a la β -tubulina describiendo dos modos de unión. En la caracterización de los dos modos de unión no se describieron interacciones con el lazo-M de la tubulina, por lo que se concluye que la dactilolida no afecta

la estructura de esta región de la proteína. Finalmente, se caracterizó el mecanismo de unión covalente de la dactilolida y sus derivados.

8 Conclusions

In this work, five alternatives to the R-NOE factor (R_t , R_{tm} , R_{twm} , R_{ms} , and R_{msw}) were designed to solve structures of molecular complexes based on Saturation Transfer Differences data. In the design of the alternatives, the terms $W_{cal,k}$, ΔM , ΔM_w , $N_i \Delta M^2$, and $N_i \Delta M_w^2$ were combined to improve the adaptation of the R-NOE factor to the molecular modeling process and to reduce its parametric dependence. Besides, the error equations of the R_e and R_t factors were obtained to evaluate the statistical significance of the results. The six R-NOE factors (R_e , R_t , R_{tm} , R_{twm} , R_{ms} , and R_{msw}) were separated into two groups according to their utilities and performances. The first group (R_{tm} , R_{twm} , R_{ms} , and R_{msw}) is the least sensitive to parametric changes. This characteristic allows R-NOE functions to locate the real structure among five candidates from a set of molecular models and independently of the chemical-physical parameters. The second group consists of R_e and R_t ; the two factors showed to be suitable for the processes of structural and parametric optimization. Although the two R-NOE factors had a similar performance, R_t was better adapted to the molecular modeling process.

Also, an automatic molecular methodology was designed based on Saturation Transfer Difference data for elucidating the structure of receptor-ligand complexes. The method was implemented in the STD-MaPa program. The STD-MaPa is composed by a friendly visual interface that allows its use for non-expert personnel.

The different parts of the methodology implemented in the STD-MaPa program were tested. The results showed that the program generates the structure describing the STD-NMR pattern 100% of the time, and identifies it using the R-NOE factors as evaluation functions. Although, results have to be obtained under certain conditions that ensure their reliability.

Based on the STD-NMR data and the coordinates of the molecular complexes, the STD-MaPa calculates the apparent association constant, the apparent rate of dissociation constant, the apparent correlation time of the ligand and the apparent correlation time of the macromolecule using a parametric adjustment algorithm. Thus, the correlation time and the degree of polymerization of the α,β -tubulin heterodimer were obtained by characterizing its binding to PM060327.

The STD-MaPa allows its adaptation to external methods for exploring 3D coordinates of receptor-ligand complexes. In our work, the conformational surfaces of the six R-NOE factors were calculated for the kifunesine-GMII, salacinol-GMII and epothilone-A- α,β -tubulin complexes by coupling the STD-MaPa with an external MonteCarlo algorithm. Results showed that epothilone-A coexists in two conformations in the binding to the heterodimer of α,β -tubulin.

On the other hand, the General Concentration Equations for CORCEMA (EGC-CORCEMA) were obtained. The equations get the CORCEMA adapted concentrations for solving structures of receptor-ligand complexes with multiple binding modes. Besides, it was found that with the STD-MaPa it is not necessary to obtain a growth curve to eliminate rebinding and T1 relaxation. Because, the program rectifies the deviation on the STD-NMR pattern and is able to solve the quantitative 3D coordinates of the molecular complexes.

The STD-MaPa was further applied to indentify the binding of derivatives of epothilone and podophyllotoxin using, as templates, the target coordinates of compounds with high structural similarities. These results were achieved coupling the STD-MaPa with various external methodologies as conformational searching and structural optimization; to which allowed adapting the conformations of the binding sites to the bind with the derivatives.

Finally, a hybrid method was designed to solve the 3D coordinates of protein-ligand complexes by combining the TR-NOESY, STD-NMR and DRX techniques. This approach allowed solving the 3D coordinates of tetramer T₂R

bound to dactylolide. Results showed that dactylolide binds to β -tubulin by describing two binding modes. Finally, the mechanism of covalent binding of dactylolide and its derivatives were further characterized.

Anexos

9 Anexos

9.1 Ecuaciones diferenciales de los factores R_t y R_e

El error de R_t y R_e se obtiene indirectamente a partir de la medida experimental de $S_{exp,k}$ y $S_{cal,k}$. La incertidumbre experimental se propaga hacia los factores R-NOE mediante el cálculo diferencial de sus variables. Esto equivale a calcular el incremento de las dos funciones a partir de sus variables, siempre y cuando, el error en estas variables sea pequeño. Además, teniendo en cuenta que los errores cometidos por cada una de las variables no pueden disminuir el error cometido por el resto de ellas, se toman los valores absolutos de las derivadas parciales.

$$dRe(S_{exp,k}, S_{cal,k}) = \sum_{k=1}^N \left| \frac{\partial Re}{\partial S_{exp,k}} \right| dS_{exp,k} + \sum_{k=1}^N \left| \frac{\partial Re}{\partial S_{cal,k}} \right| dS_{cal,k} \quad 9.1.1$$

$$dRt(S_{exp,k}, S_{cal,k}) = \sum_{k=1}^N \left| \frac{\partial Rt}{\partial S_{exp,k}} \right| dS_{exp,k} + \sum_{k=1}^N \left| \frac{\partial Rt}{\partial S_{cal,k}} \right| dS_{cal,k} \quad 9.1.2$$

Los valores de $S_{cal,k}$ se obtienen a partir de la resolución de la ecuación del CORCEMA mediante el ajuste paramétrico de diversas constantes químico-físicas. Este procedimiento es analítico y la incertidumbre está dada por el truncamiento numérico que se realiza en el cálculo. Los algoritmos modernos incluyen variables de doble precisión, lo que permite describir los números hasta la decimosexta cifra decimal. Dado que el error experimental se ubica en torno a la primera cifra, el error de $dS_{cal,k}$ se considera despreciable.

Cuando $k=p$, siendo p un número natural definido en el intervalo $1 \leq p \leq N$, nos quedan las derivadas parciales como $\frac{\partial R_e}{\partial S_{exp,p}}$ y $\frac{\partial R_t}{\partial S_{cal,p}}$. Para resolver la derivada parcial cuando $k=p$ aplicamos las reglas de derivabilidad.

$$\begin{aligned} & \frac{\partial R_e}{\partial S_{exp,p}} \\ &= \frac{1}{R_e} \left\{ \frac{\left[W_{exp,p}(S_{exp,p} - S_{cal,p}) - \frac{1}{2} W_{exp,p}^2 (S_{exp,p} - S_{cal,p})^2 \right] \sum_{k=1}^N W_{exp,k} S_{exp,k}^2}{\left[\sum_{k=1}^N W_{exp,k} S_{exp,k}^2 \right] \left[\sum_{k=1}^N W_{exp,k} S_{exp,k}^2 \right]} \right. \\ & \quad \left. - \frac{\left[W_{exp,p} S_{exp,p} - \frac{1}{2} W_{exp,p}^2 S_{exp,p}^2 \right] \sum_{k=1}^N W_{exp,k} (S_{exp,k} - S_{cal,k})^2}{\left[\sum_{k=1}^N W_{exp,k} S_{exp,k}^2 \right] \left[\sum_{k=1}^N W_{exp,k} S_{exp,k}^2 \right]} \right\} \end{aligned} \quad 9.1.3$$

$$\begin{aligned} \frac{\partial R_t}{\partial S_{exp,p}} &= \frac{1}{R_t} \left\{ \frac{W_{cal,p}(S_{exp,p} - S_{cal,p}) \sum_{k=1}^N W_{cal,k} S_{exp,k}^2}{\left[\sum_{k=1}^N W_{cal,k} S_{exp,k}^2 \right] \left[\sum_{k=1}^N W_{cal,k} S_{exp,k}^2 \right]} \right. \\ & \quad \left. - \frac{W_{cal,p} S_{exp,p} \sum_{k=1}^N W_{cal,k} (S_{exp,k} - S_{cal,k})^2}{\left[\sum_{k=1}^N W_{cal,k} S_{exp,k}^2 \right] \left[\sum_{k=1}^N W_{cal,k} S_{exp,k}^2 \right]} \right\} \end{aligned} \quad 9.1.4$$

Los términos de las ecuaciones resultantes son simplificados mediante la adición de los factores R_e y R_t .

$$\begin{aligned} \frac{\partial R_e}{\partial S_{exp,p}} &= R_e \left\{ \frac{W_{exp,p}(S_{exp,p} - S_{cal,p})}{\sum_{k=1}^N W_{exp,k} (S_{exp,k} - S_{cal,k})^2} - \frac{W_{exp,p} S_{exp,p}}{\sum_{k=1}^N W_{exp,k} S_{exp,k}^2} \right. \\ & \quad \left. - \frac{1}{2} \left[\frac{W_{exp,p}^2 (S_{exp,p} - S_{cal,p})^2}{\sum_{k=1}^N W_{exp,k} (S_{exp,k} - S_{cal,k})^2} - \frac{W_{exp,p}^2 S_{exp,p}^2}{\sum_{k=1}^N W_{exp,k} S_{exp,k}^2} \right] \right\} \end{aligned} \quad 9.1.5$$

$$\frac{\partial R_t}{\partial S_{exp,p}} = R_t \left\{ \frac{W_{cal,p}(S_{exp,p} - S_{cal,p})}{\sum_{k=1}^N W_{cal,k}(S_{exp,k} - S_{cal,k})^2} - \frac{W_{cal,p}S_{exp,p}}{\sum_{k=1}^N W_{cal,k}S_{exp,k}^2} \right\} \quad 9.1.6$$

Si $N=1$, las expresiones de 9.1.5 y 9.1.6 se igualan, dado que el segundo término de la derivada parcial de R_e para $k=p$ es cero. En este caso, el error relativo de ambas ecuaciones depende de la diferencia entre la inversa de la sustracción de las intensidades y del factor de peso experimental $\left(\frac{1}{S_{exp,k} - S_{cal,k}} - \frac{1}{S_{exp,k}} \right)$. Finalmente, las ecuaciones 9.1.5 y 9.1.6 son sustituidas en 9.1.1 y 9.1.2 para determinar los errores de R_e y R_t .

9.2 Pruebas estadísticas de Levene, Barlett y Shapiro-Wilk

En este anexo se va a comprobar si las variaciones máximas normalizadas de los factores R-NOE cumplen con las condiciones necesarias para poder aplicar la prueba estadística de la ANOVA y el test de Tukey. Nuestro ensayo estadístico está dividido en seis grupos diferentes, uno por cada factor R-NOE. El estudio se realiza a partir de los datos extraídos de los complejos de kifunesina-dGMII, salacino-dGMII y epotilona-A- β -tubulina. Por tanto, el tamaño poblacional de cada grupo es igual a tres. En nuestro análisis, la independencia de los datos es asumible porque los cálculos del factor R-NOE y los STD-NMR fueron realizados de forma independiente. Sin embargo, la normalidad de las poblaciones y la igualdad entre varianzas se va a comprobar mediante tres pruebas estadísticas (Ver Figura 9.2.1).

Primero, y dada su importancia, se comprobó la homocedasticidad muestral mediante las pruebas estadísticas de Levene y Barlett. Se eligieron éstas dos porque difieren en la respuesta a la desviación de la normalidad. Para ambas pruebas, hemos tenido en cuenta la hipótesis nula, en la que se considera que las varianzas entre todas las poblaciones son iguales y un error asumido del 5%. Bajo estas condiciones, las pruebas estadísticas de Levene y Barlett concordaron en que todas las muestras cumplen con la hipótesis nula, dando un p-valor > 0.5 . Por esta razón, se concluye que las varianzas muestrales son similares.

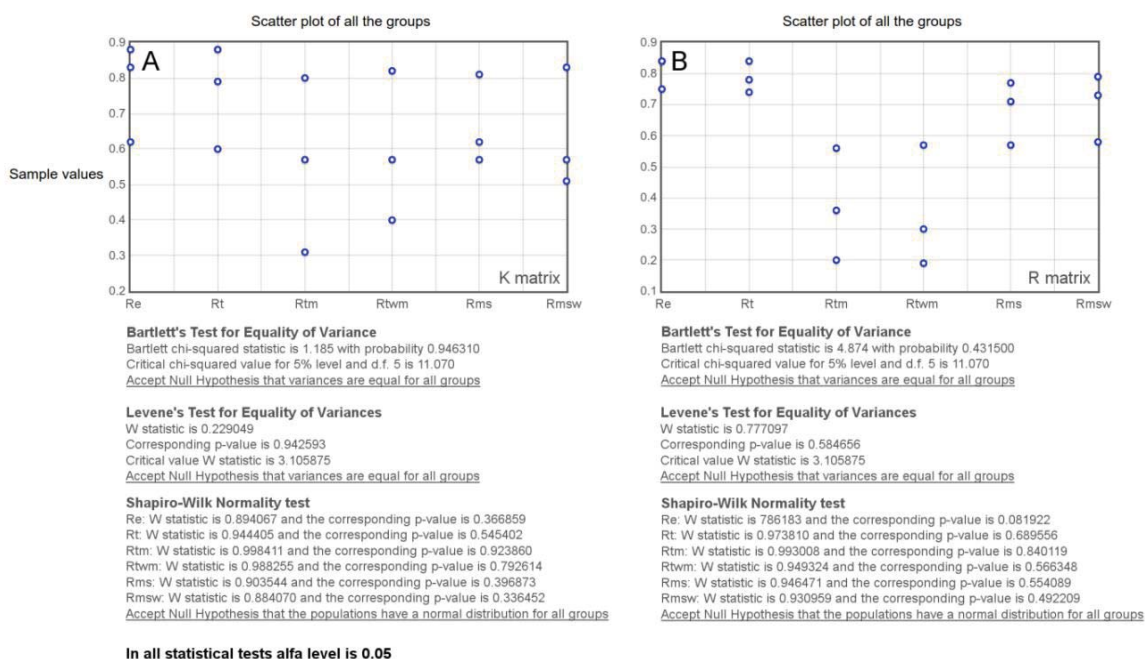


Figura 9.2.1. En la parte superior, se representan en los gráficos los valores máximos de las desviaciones de los factores R-NOE, al variar los parámetros químico-físicos para los tres complejos moleculares. En la parte inferior aparece el estudio estadístico de igualdad de varianza de Levene (<http://scistatcalc.blogspot.com.es/2013/12/levenes-test-equality-of-variances.html>), el de Barlett (<http://scistatcalc.blogspot.com.es/2013/10/bartletts-test-calculator.html>) y la prueba de normalidad Shapiro-Wilk realizada con el programa QtiPlot. En la izquierda se encuentran las variaciones máximas de los factores R-NOE para el sistema con kifunesina, salacinol y epotilona-A al modificar los parámetros de la matriz cinética (**K**). En la izquierda se encuentran las variaciones máximas de los factores R-NOE para el sistema con kifunesina, salacinol y epotilona-A al modificar los parámetros de la matriz de relajación (**R**).

El mismo procedimiento se ha realizado para el estudio de la normalidad muestral. En esa parte, se ha utilizado la prueba de Shapiro-Wilk para evaluar si las muestras tienen una distribución normal. Ahí también se asumió un error

del 5%. En todos los casos, el p-valor es mayor que 0.5, por lo que se puede considerar que todas las poblaciones presentan una distribución poblacional normal.

9.3 Publicaciones

- 1- Yong Tang, Javier Rodríguez-Salarichs, Yu Zhao, Pei Cai, Juan Estévez-Gallego, Francisco Balaguer-Pérez, Mariano Redondo Horcajo, Daniel Lucena-Agell, Isabel Barasoain, J Fernando Díaz, Wei-Shuo Fang. Modification of C-seco taxoids through ring tethering and substituent replacement leading to effective agents against tumor drug resistance mediated by β III-Tubulin and P-glycoprotein (P-gp) overexpressions. European Journal of Medicinal Chemistry. 03/06/2017. DOI 10.1016/j.ejmech.2017.06.001.

Tipo de producción: Artículo

Tipo de soporte: Revista

Fuente de citas: European Journal of Medicinal Chemistry **Citas:** 0

- 2- Saez-Calvo, Gonzalo; Sharma, Ashwani; Balaguer, Francisco de Asis; Barasoain, Isabel; **Rodríguez-Salarichs, Javier**; Olieric, Natacha; Munoz-Hernandez, Hugo; Berbis, Manuel Alvaro; Wendeborn, Sebastian; Penalva, Miguel Angel; Matesanz, Ruth; Canales, Angeles; Prota, Andrea Enrico; Jimenez-Barbero, Jesus; Andreu, Jose Manuel; Lamberth, Clemens; Steinmetz, Michel Olivier; Diaz, Jose Fernando. Triazolopyrimidines Are Microtubule-Stabilizing Agents that Bind the Vinca Inhibitor Site of Tubulin. Cell chemical biology. 31/05/2017. ISSN 2451-9456

Tipo de producción: Artículo

Tipo de soporte: Revista

Fuente de citas: WOS

Citas: 0

- 3- Jantsch, Andrea; Nieto, Lidia; Gertsch, Jurg; **Rodríguez-Salarichs, Javier**; Matesanz, Ruth; Jimenez-Barbero, Jesus; Fernando Diaz, J.; Canales, Angeles; Altmann, Karl-Heinz. Synthesis, Biological Profiling and Determination of the Tubulin-Bound Conformation of 12-Aza-

Epothilones (Azathilones). MOLECULES. 21 - 8, MDPI AG, 01/08/2016.
ISSN 1420-3049

Tipo de producción: Artículo

Tipo de soporte: Revista

Fuente de citas: WOS

Citas: 0

- 4- Antunez-Mojica, Mayra; Rodriguez-Salarichs, Javier; Redondo-Horcajo, Mariano; Leon, Alejandra; Barasoain, Isabel; Canales, Angeles; Canada, F. J.; Jimenez-Barbero, Jesus; Alvarez, Laura; Fernando Diaz, J. Structural and Biochemical Characterization of the Interaction of Tubulin with Potent Natural Analogues of Podophyllotoxin. JOURNAL OF NATURAL PRODUCTS. 79 - 8, pp. 2113 - 2121. AMER CHEMICAL SOC, 01/08/2016. ISSN 0163-3864, ISSN 1520-6025

Tipo de producción: Artículo

Tipo de soporte: Revista

Fuente de citas: WOS

Citas: 0

- 5- Pinar, Mario; Arst, Herbert N., Jr.; Pantazopoulou, Areti; Tagua, Victor G.; de los Rios, Vivian; Rodriguez-Salarichs, Javier; Fernando Diaz, J.; Penalva, Miguel A. TRAPP II regulates exocytic Golgi exit by mediating nucleotide exchange on the Ypt31 ortholog RabE(RAB11). PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA. 112 - 14, pp. 4346 - 4351. NATL ACAD SCIENCES, 07/04/2015. ISSN 0027-8424

Tipo de producción: Artículo

Tipo de soporte: Revista

Fuente de impacto: WOS (JCR)

Categoría: Science Edition – MULTIDISCIPLINARY SCIENCES

Índice de impacto: 9.423

Revista dentro del 25%: Si

Posición de publicación: 4

Num. revistas en cat.: 63

Fuente de citas: WOS

Citas: 13

- 6- Matesanz, Ruth; Trigili, Chiara; **Rodriguez-Salarichs, Javier**; Zanardi, Ilaria; Pera, Benet; Nogales, Aurora; Fang, Wei-Shuo; Jimenez-Barbero, Jesus; Canales, Angeles; Barasoain, Isabel; Ojima, Iwao; Fernando Diaz, J. Taxanes with high potency inducing tubulin assembly overcome tumoural cell resistances. BIOORGANIC & MEDICINAL CHEMISTRY. 22 - 18, pp. 5078 - 5090. PERGAMON-ELSEVIER SCIENCE LTD, 15/09/2014. ISSN 0968-0896, ISSN 1464-3391

Tipo de producción: Artículo

Tipo de soporte: Revista

Fuente de impacto: WOS (JCR)

Categoría: Science Edition - BIOCHEMISTRY & MOLECULAR BIOLOGY

Índice de impacto: 2.793

Revista dentro del 25%: No

Posición de publicación: 140

Num. revistas en cat.: 290

Fuente de impacto: WOS (JCR)

Categoría: Science Edition - CHEMISTRY, MEDICINAL

Índice de impacto: 2.793

Revista dentro del 25%: No

Posición de publicación: 24

Num. revistas en cat.: 59

Fuente de impacto: WOS (JCR) **Categoría:** Science Edition - CHEMISTRY, ORGANIC

Índice de impacto: 2.793

Revista dentro del 25%: No

Posición de publicación: 15

Num. revistas en cat.: 58

Fuente de citas: WOS

Citas: 6

- 7- Canales, Angeles; Nieto, Lidia; **Rodriguez-Salarichs, Javier**; Sanchez-Murcia, Pedro A.; Coderch, Claire; Cortes-Cabrera, Alvaro; Paterson, Ian; Carlomagno, Teresa; Gago, Federico; Andreu, Jose M.; Altmann, Karl-Heinz; Jimenez-Barbero, Jesus; Fernando Diaz, J. Molecular Recognition of Epothilones by Microtubules and Tubulin Dimers Revealed by Biochemical and NMR Approaches. ACS CHEMICAL BIOLOGY. 9 - 4, pp. 1033 - 1043. AMER CHEMICAL SOC, 01/04/2014. ISSN 1554-8929, ISSN 1554-8937

Tipo de producción: Artículo

Tipo de soporte: Revista

Fuente de impacto: WOS (JCR)

Categoría: Science Edition - BIOCHEMISTRY & MOLECULAR BIOLOGY

Índice de impacto: 5.331

Revista dentro del 25%: Si

Posición de publicación: 46

Num. revistas en cat.: 290

Fuente de citas: WOS

Citas: 10

- 8- Pera, Benet; Barasoain, Isabel; Pantazopoulou, Areti; Canales, Angeles; Matesanz, Ruth; **Rodriguez-Salarichs, Javier**; Garcia-Fernandez, Luis F.; Moneo, Victoria; Jimenez-Barbero, Jesus; Galmarini, Carlos M.; Cuevas, Carmen; Penalva, Miguel A.; Fernando Diaz, J.; Andreu, Jose M. New Interfacial Microtubule Inhibitors of Marine Origin, PM050489/PM060184, with Potent Antitumor Activity and a Distinct Mechanism. ACS CHEMICAL BIOLOGY. 8 - 9, pp. 2084 - 2094. AMER CHEMICAL SOC, 01/09/2013. ISSN 1554-8929, ISSN 1554-8937

Tipo de producción: Artículo

Tipo de soporte: Revista

Fuente de impacto: WOS (JCR)

Categoría: Science Edition - BIOCHEMISTRY & MOLECULAR
BIOLOGY

Índice de impacto: 5.356

Revista dentro del 25%: Si

Posición de publicación: 51

Num. revistas en cat.: 291

Fuente de citas: WOS

Citas: 8

- 9- Maccari, Giorgio; Mori, Mattia; Rodriguez-Salarichs, Javier; Fang, Weishuo; Fernando Diaz, Jose; Botta, Maurizio. Free Energy Profile and Kinetics Studies of Paclitaxel Internalization from the Outer to the Inner Wall of Microtubules. JOURNAL OF CHEMICAL THEORY AND COMPUTATION. 9 - 1, pp. 698 - 706. AMER CHEMICAL SOC, 01/01/2013. ISSN 1549-9618, ISSN 1549-9626

Tipo de producción: Artículo

Tipo de soporte: Revista

Fuente de impacto: WOS (JCR)

Categoría: Science Edition - CHEMISTRY, PHYSICAL

Índice de impacto: 5.31

Revista dentro del 25%: Si

Posición de publicación: 25

Num. revistas en cat.: 136

Fuente de impacto: WOS (JCR)

Categoría: Science Edition - PHYSICS, ATOMIC, MOLECULAR &
CHEMICAL

Índice de impacto: 5.31

Revista dentro del 25%: Si

Posición de publicación: 3

Num. revistas en cat.: 33

Fuente de citas: WOS

Citas: 4

10-Lam-Barandela, A.; Rodríguez-Salarichs, J.; Rojas-Lorenzo, G. A.; Álvarez-Noguera, L. J.; Rivera-Denis, A. Aggregation Process of a Benzyl Ammonium Chloride: Molecular Dynamics Simulation. REVISTA CUBANA DE QUÍMICA, 25 - 1, pp. 66 - 74. 01/01/2013.

Tipo de producción: Artículo

Tipo de soporte: Revista

11-Field, Jessica J.; Pera, Benet; Calvo, Enrique; Canales, Angeles; Zurwerra, Didier; Trigili, Chiara; Rodriguez-Salarichs, Javier; Matesanz, Ruth; Kanakkanthara, Arun; Wakefield, St John; Singh, A. Jonathan; Jimenez-Barbero, Jesus; Northcote, Peter; Miller, John H.; Antonio Lopez, Juan; Hamel, Ernest; Barasoain, Isabel; Altmann, Karl-Heinz; Fernando Diaz, Jose. Zampanolide, a Potent New Microtubule-Stabilizing Agent, Covalently Reacts with the Taxane Luminal Site in Tubulin alpha,beta-Heterodimers and Microtubules. CHEMISTRY & BIOLOGY. 19 - 6, pp. 686 - 698. CELL PRESS, 22/06/2012. ISSN 1074-5521

Tipo de producción: Artículo

Tipo de soporte: Revista

Fuente de impacto: WOS (JCR)

Categoría: Science Edition - BIOCHEMISTRY & MOLECULAR BIOLOGY

Índice de impacto: 6.157

Revista dentro del 25%: Si

Posición de publicación: 39

Num. revistas en cat.: 290

Fuente de citas: WOS

Citas: 32

12-Matesanz, Ruth; **Rodriguez-Salarichs, Javier**; Pera, Benet; Canales, Angeles; Manuel Andreu, Jose; Jimenez-Barbero, Jesus; Bras, Wim; Nogales, Aurora; Fang, Wei-Shuo; Fernando Diaz, Jose. Modulation of Microtubule Interprotofilament Interactions by Modified Taxanes. BIOPHYSICAL JOURNAL. 101 - 12, pp. 2970 - 2980. CELL PRESS, 21/12/2011. ISSN 0006-3495

Tipo de producción: Artículo

Tipo de soporte: Revista

Fuente de impacto: WOS (JCR)

Categoría: Science Edition - BIOPHYSICS

Índice de impacto: 3.653

Revista dentro del 25%: No

Posición de publicación: 19

Num. revistas en cat.: 74

Fuente de citas: WOS

Citas: 17

13-Canales, Angeles; **Rodriguez-Salarichs, Javier**; Trigili, Chiara; Nieto, Lidia; Coderch, Claire; Andreu, Jose Manuel; Paterson, Ian; Jimenez-Barbero, Jesus; Fernando Diaz, J. Insights into the interaction of Discodermolide and Docetaxel with Tubulin. Mapping the Binding Sites of Microtubule-Stabilizing Agents by Using an Integrated NMR and Computational Approach. ACS CHEMICAL BIOLOGY. 6 - 8, pp. 789 - 799. AMER CHEMICAL SOC, 01/08/2011. ISSN 1554-8929

Tipo de producción: Artículo

Tipo de soporte: Revista

Fuente de impacto: WOS (JCR)

Categoría: Science Edition - BIOCHEMISTRY & MOLECULAR BIOLOGY

Índice de impacto: 6.446

Revista dentro del 25%: Si

Posición de publicación: 36

Num. revistas en cat.: 290

Fuente de citas: WOS

Citas: 31

14-Rodriguez, Hortensia; Coro, Julieta; Lam, Anabel; Salfran, Esperanza; **Rodriguez-Salarichs, Javier**; Suarez, Margarita; Albericio, Fernando; Martin, Nazario. High-throughput preparation of alkyl 4-aryl substituted-2-methyl-6-thioxo-1,4,5,6-tetrahydropyridine-3-carboxylates under microwave irradiation. ARKIVOC. pp. 125 - U1. ARKAT USA INC, 01/01/2011. ISSN 1551-7004

Tipo de producción: Artículo

Tipo de soporte: Revista

Fuente de impacto: WOS (JCR)

Categoría: Science Edition - CHEMISTRY, ORGANIC

Índice de impacto: 1.252

Revista dentro del 25%: No

Posición de publicación: 34

Num. revistas en cat.: 56

Fuente de citas: WOS

Citas: 3

Bibliografía

10 Bibliografía

Adams, P. D., P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart (2010). "PHENIX: a comprehensive Python-based system for macromolecular structure solution." Acta Crystallographica Section D **66**(2): 213-221.

Allen, M. P. and D. J. Tildesley (1989). Computer simulation of liquids, Oxford university press.

Andreu, J. M. (2007). Tubulin Purification. Methods in Molecular Medicine. Totowa, NJ, Humana Press Inc. **137**: 17-28.

Andreu, J. M., M. J. Gorbunoff, J. C. Lee and S. N. Timasheff (1984). "Interaction of tubulin with bifunctional colchicine analogues: an equilibrium study." Biochemistry **23**: 1742-1752.

Andreu, J. M., M. a. A. Oliva and O. Monasterio (2002). "Reversible Unfolding of FtsZ Cell Division Proteins from Archaea and Bacteria COMPARISON WITH EUKARYOTIC TUBULIN FOLDING AND ASSEMBLY." Journal of Biological Chemistry **277**(45): 43262-43270.

Angulo, J., I. Díaz, J. J. Reina, G. Tabarani, F. Fieschi, J. Rojo and P. M. Nieto (2008). "Saturation Transfer Difference (STD) NMR Spectroscopy Characterization of Dual Binding Mode of a Mannose Disaccharide to DC-SIGN." ChemBioChem **9**(14): 2225-2227.

Angulo, J., P. M. Enríquez-Navas and P. M. Nieto (2010). "Ligand–Receptor Binding Affinities from Saturation Transfer Difference (STD) NMR Spectroscopy: The Binding Isotherm of STD Initial Growth Rates." Chemistry – A European Journal **16**(26): 7803-7812.

Angulo, J. and P. M. Nieto (2011). Espectroscopía RMN de Diferencia de Transferencia de Saturación (STD NMR). Aplicaciones en Química del Reconocimiento Molecular.

Angulo, J. and P. M. Nieto (2011). "STD-NMR: application to transient interactions between biomolecules—a quantitative approach." European Biophysics Journal **40**(12): 1357-1369.

Antúñez-Mojica, M., J. Rodríguez-Salarichs, M. Redondo-Horcajo, A. León, I. Barasoain, A. Canales, F. J. Cañada, J. Jiménez-Barbero, L. Alvarez and J. F. Díaz (2016). "Structural and Biochemical Characterization of the Interaction of Tubulin with Potent Natural Analogues of Podophyllotoxin." Journal of Natural Products **79**(8): 2113-2121.

Assadi-Porter, F. M., M. Tonelli, E. Maillet, K. Hallenga, O. Benard, M. Max and J. L. Markley (2008). "Direct NMR Detection of the Binding of Functional Ligands to the Human Sweet Receptor, a Heterodimeric Family 3 GPCR." Journal of the American Chemical Society **130**(23): 7212-7213.

Balaram, P., A. A. Bothner-By and E. Breslow (1972). "Localization of tyrosine at the binding site of neurophysin II by negative nuclear Overhauser effects." Journal of the American Chemical Society **94**(11): 4017-4018.

Balaram, P., A. A. Bothner-By and J. Dadok (1972). "Negative nuclear Overhauser effects as probes of macromolecular structure." Journal of the American Chemical Society **94**(11): 4015-4017.

Barra, P. A., V. A. Jiménez, J. A. Gavin, A. H. Daranas and J. B. Alderete (2016). "Discovery of New E-Selectin Inhibitors by Virtual Screening, Fluorescence Binding Assays, and STD NMR Experiments." ChemMedChem **11**(9): 1008-1014.

Bax, A. (1989). "Two-Dimensional NMR and Protein Structure." Annual Review of Biochemistry **58**(1): 223-256.

Bax, A. and S. Grzesiek (1993). "Methodological advances in protein NMR." Accounts of Chemical Research **26**(4): 131-138.

Baxter, J. (1981). "Local Optima Avoidance in Depot Location." The Journal of the Operational Research Society **32**(9): 815-819.

Bernardi, A., D. Potenza, A. M. Capelli, A. García-Herrero, F. J. Cañada and J. Jiménez-Barbero (2002). "Second-generation mimics of ganglioside GM1 oligosaccharide: A three-dimensional view of their interactions with bacterial enterotoxins by NMR and computational methods." Chemistry-a European Journal **8**: 4598-4612.

Bhalla, K. N. (2003). "Microtubule-targeted anticancer agents and apoptosis." Oncogene **22**: 9075-9086.

Birdsall, B., J. Feeney, G. C. K. Roberts and A. S. V. Burgen (1980). "The use of saturation transfer NMR experiments to monitor the conformational selection accompanying ligand—protein interactions." FEBS Letters **120**(1): 107-109.

Bjellmar, P., P. Larsson, M. A. Cuendet, B. Hess and E. Lindahl (2010). "Implementation of the CHARMM Force Field in GROMACS: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models." Journal of Chemical Theory and Computation **6**(2): 459-466.

Bloch, F. (1946). "Nuclear Induction." Physical Review **70**(7-8): 460-474.

Blum, C., A. Roli and M. Sampels (2008). Hybrid Metaheuristics: An Emerging Approach to Optimization, Springer Berlin Heidelberg.

Bragg, W. H. and W. L. Bragg (1913). "The Reflection of X-rays by Crystals." Proceedings of the Royal Society of London. Series A **88**(605): 428.

Brüschweiler, R. and D. A. Case (1994). "Characterization of biomolecular structure and dynamics by NMR cross relaxation." Progress in nuclear magnetic resonance spectroscopy **26**: 27-58.

Buey, R. M., I. Barasoain, E. Jackson, A. Meyer, P. Giannakakou, I. Paterson, S. Mooberry, J. M. Andreu and J. F. Díaz (2005). "Microtubule interactions with chemically diverse stabilizing agents: Thermodynamics of binding to the paclitaxel site predicts cytotoxicity." Chem Biol **12**: 1269-1279.

Buey, R. M., E. Calvo, I. Barasoain, O. Pineda, M. C. Edler, R. Matesanz, G. Cerezo, C. D. Vanderwal, B. W. Day, E. J. Sorensen, J. A. Lopez, J. M. Andreu, E. Hamel and J. F. Díaz (2007). "Cyclostreptin binds covalently to microtubule pores and lumenal taxoid binding sites." Nature Chem Biol **3**: 117-125.

Bursulaya, B. D., M. Totrov, R. Abagyan and C. L. Brooks (2003). "Comparative study of several algorithms for flexible ligand docking." Journal of computer-aided molecular design **17**(11): 755-763.

Canales, A., R. Matesanz, N. M. Gardner, J. M. Andreu, I. Paterson, J. F. Diaz and J. Jimenez-Barbero (2008). "The Bound Conformation of Microtubule-Stabilizing Agents: NMR Insights into the Bioactive 3D Structure of Discodermolide and Dictyostatin." Chemistry Eur. J. **14**: 7557-7569.

Canales, A., L. Nieto, J. Rodríguez-Salarichs, P. A. Sánchez-Murcia, C. Coderch, A. C. Cabrera, I. Paterson, T. Carlomagno, F. Gago and J. M. Andreu (2014). "The molecular recognition of epothilones by microtubules and tubulin dimers revealed by biochemical and NMR approaches." ACS chemical biology.

Canales, A., J. Rodríguez-Salarichs, C. Trigili, L. Nieto, C. Coderch, J. M. Andreu, I. Paterson, J. Jiménez-Barbero and J. F. Díaz (2011). "Insights into the interaction of discodermolide and docetaxel with tubulin. Mapping the binding sites of microtubule-stabilizing agents by using an integrated NMR and computational approach." ACS chemical biology **6**(8): 789-799.

Canales, A., J. R. Salarichs, C. Trigili, L. Nieto, C. Coderch, J. M. Andreu, I. Paterson, J. Jiménez-Barbero and J. F. Díaz (2011). "Insights into the interaction of discodermolide and docetaxel with dimeric tubulin. Mapping the

binding sites of microtubule-stabilizing agents using an integrated NMR and computational approach." ACS Chemical Biology **6**: 789-799.

Carver, T. R. and C. P. Slichter (1953). "Polarization of nuclear spins in metals." Physical Review **92**(1): 212.

Casewit, C. J., K. S. Colwell and A. K. Rappe (1992). "Application of a universal force field to organic molecules." Journal of the American Chemical Society **114**(25): 10035-10046.

Cavanagh, J., W. J. Fairbrother, A. G. Palmer, N. J. Skelton and M. Rance (2010). Protein NMR Spectroscopy: Principles and Practice, Elsevier Science.

Cayley, P. J., J. P. Albrand, J. Feeney, G. C. K. Roberts, E. A. Piper and A. S. V. Burgen (1979). "Nuclear magnetic resonance studies of the binding of trimethoprim to dihydrofolate reductase." Biochemistry **18**(18): 3886-3895.

Cieplak, P., W. D. Cornell, C. Bayly and P. A. Kollman (1995). "Application of the multimolecule and multiconformational resp methodology to biopolymers - charge derivation for DNA, RNA, and proteins." J. Comput. Chem. **16**: 1357-1377.

Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman (1995). "A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules." J Am Chem Soc **117**: 5179-5197.

Cutignano, A., I. Bruno, G. Bifulco, A. Casapullo, C. Debitus, L. Gomez-Paloma and R. Riccio (2001). "Dactylolide, a New Cytotoxic Macrolide from the Vanuatu Sponge *Dactylospongia* sp." European Journal of Organic Chemistry **2001**: 775-778.

Chang, G., W. C. Guida and W. C. Still (1989). "An internal-coordinate Monte Carlo method for searching conformational space." Journal of the American Chemical Society **111**(12): 4379-4386.

Checchi, P. M., J. H. Nettles, J. Zhou, J. P. Snyder and H. C. Joshi (2003). "Microtubule-interacting drugs for cancer treatment." Trends in Pharmacological Sciences **24**(7): 361-365.

Cheshnovsky, D. and G. Navon (1980). "Nuclear magnetic resonance studies of carbonic anhydrase catalyzed reversible hydration of acetaldehyde by the saturation transfer method." Biochemistry **19**(9): 1866-1873.

Cho, C. H., J. Urquidi, S. Singh and G. W. Robinson (1999). "Thermal Offset Viscosities of Liquid H₂O, D₂O, and T₂O." The Journal of Physical Chemistry B **103**(11): 1991-1994.

Davis, I. W., L. W. Murray, J. S. Richardson and D. C. Richardson (2004). "MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes." Nucleic acids research **32**(suppl 2): W615-W619.

Dellwo, M. J. and A. J. Wand (1993). "The influence of methyl rotor dynamics on hydrogen relaxation networks: derivation of spectral densities in model-free form." Journal of the American Chemical Society **115**(5): 1886-1893.

Díaz, J. F. and J. M. Andreu (1993). "Assembly of purified GDP-tubulin into microtubules induced by taxol and taxotere: reversibility, ligand stoichiometry, and competition." Biochemistry **32**: 2747-2755.

Díaz, J. F., I. Barasoain and J. M. Andreu (2003). "Fast kinetics of Taxol binding to microtubules. Effects of solution variables and microtubule-associated proteins." J Biol Chem **278**: 8407-8419.

Díaz, J. F., E. Pantos, J. Bordas and J. M. Andreu (1994). "Solution Structure of GDP-tubulin Double Rings to 3 nm Resolution and Comparison with Microtubules." Journal of Molecular Biology **238**(2): 214-225.

Drenth, J. (2013). Principles of Protein X-ray Crystallography, Springer New York.

Edler, M. C., R. M. Buey, R. Gussio, A. I. Marcus, C. D. Vanderwal, E. J. Sorensen, J. F. Diaz, P. Giannakakou and E. Hamel (2005). "Cyclostreptin (FR182877), an Antitumor Tubulin-Polymerizing Agent Deficient in Enhancing Tubulin Assembly Despite Its High Affinity for the Taxoid Site." Biochemistry **44**: 11525-11538.

Enríquez-Navas, P. M., M. Marradi, D. Padro, J. Angulo and S. Penadés (2011). "A Solution NMR Study of the Interactions of Oligomannosides and the Anti-HIV-1 2G12 Antibody Reveals Distinct Binding Modes for Branched Ligands." Chemistry–A European Journal **17**(5): 1547-1560.

Ernst, R. R. (1966). "Nuclear Magnetic Double Resonance with an Incoherent Radio-Frequency Field." The Journal of Chemical Physics **45**(10): 3845-3861.

Ernst, R. R. and W. A. Anderson (1966). "Application of Fourier transform spectroscopy to magnetic resonance." Review of Scientific Instruments **37**(1): 93-102.

Field, J. J., B. Pera, E. Calvo, A. Canales, D. Zurwerra, C. Trigili, J. Rodríguez-Salarichs, R. Matesanz, A. Kanakkanthara and S. J. Wakefield (2012). "Zampanolide, a potent new microtubule-stabilizing agent, covalently reacts with the taxane luminal site in tubulin α , β -heterodimers and microtubules." Chemistry & biology **19**(6): 686-698.

Garcia de la Torre, J., M. L. Huertas and B. Carrasco (2000). "Calculation of hydrodynamic properties of globular proteins from their atomic-level structure." Biophysical Journal **78**: 719-730.

Garcia de la Torre, J., S. Navarro, M. C. Lopez Martinez, F. G. Diaz and J. J. Lopez Cascales (1994). "HYDRO: a computer program for the prediction of hydrodynamic properties of macromolecules." Biophys J **67**: 530-531.

Gigant, B., C. Wang, R. B. Ravelli, F. Roussi, M. O. Steinmetz, P. A. Curmi, A. Sobel and M. Knossow (2005). "Structural basis for the regulation of tubulin by vinblastine." Nature **435**: 519-522.

Goldman, M. (1994). NMR and more in honour of Anatole Abragam. France, Editions de la Physique.

Gonzalez, C., J. A. C. Rullmann, A. Bonvin, R. Boelens and R. Kaptein (1991). "Toward an NMR R factor." Journal of magnetic resonance **91**(3): 659-664.

Grootenhuis, P. D. J. and C. A. G. Haasnoot (1993). "A CHARMM Based Force Field for Carbohydrates Using the CHEAT Approach: Carbohydrate Hydroxyl Groups Represented by Extended Atoms." Molecular Simulation **10**(2-6): 75-95.

Haselhorst, T., H. Blanchard, M. Frank, M. J. Kraschnefski, M. J. Kiefel, A. J. Szyzew, J. C. Dyason, F. Fleming, G. Holloway, B. S. Coulson and M. von Itzstein (2007). "STD NMR spectroscopy and molecular modeling investigation of the binding of N-acetylneuraminic acid derivatives to rhesus rotavirus VP8* core (vol 17, pg 68, 2007)." Glycobiology **17**(10): 1030-1030.

Hastings, W. K. (1970). "Monte Carlo sampling methods using Markov chains and their applications." Biometrika **57**(1): 97-109.

Hatcher, E. R., O. Guvench and A. D. MacKerell (2009). "CHARMM Additive All-Atom Force Field for Acyclic Polyalcohols, Acyclic Carbohydrates, and Inositol." Journal of Chemical Theory and Computation **5**(5): 1315-1327.

Havel, T. and K. Wüthrich (1984). "A distance geometry program for determining the structures of small proteins and other macromolecules from nuclear magnetic resonance measurements of intramolecular ^1H - ^1H proximities in solution." Bulletin of mathematical biology **46**(4): 673-698.

Hayafune, M., R. Berisio, R. Marchetti, A. Silipo, M. Kayama, Y. Desaki, S. Arima, F. Squeglia, A. Ruggiero, K. Tokuyasu, A. Molinaro, H. Kaku and N. Shibuya (2014). "Chitin-induced activation of immune signaling by the rice

receptor CEBiP relies on a unique sandwich-type dimerization." Proceedings of the National Academy of Sciences **111**(3): E404-E413.

Hendrickson, W. A. (1986). "The 1985 Nobel Prize in Chemistry." Science **231**(4736): 362.

Hetenyi, C. and D. van der Spoel (2002). "Efficient docking of peptides to proteins without prior knowledge of the binding site." Prot. Sci. **11**: 1729-1737.

Holland, J. H. (1975). "Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence."

Honnappa, S., B. Cutting, W. Jahnke, J. Seelig and M. O. Steinmetz (2003). "Thermodynamics of the Op18/stathmin-tubulin interaction." Journal of Biological Chemistry **278**(40): 38926-38934.

Honore, S., K. Kamath, D. Braguer, L. Wilson, C. Briand and M. A. Jordan (2003). "Suppression of microtubule dynamics by discodermolide by a novel mechanism is associated with mitotic arrest and inhibition of tumor cell proliferation." Mol Cancer Ther **2**: 1303-1311.

Huang, N., B. K. Shoichet and J. J. Irwin (2006). "Benchmarking Sets for Molecular Docking." Journal of Medicinal Chemistry **49**(23): 6789-6801.

Huey, R., G. M. Morris, A. J. Olson and D. S. Goodsell (2007). "A semiempirical free energy force field with charge-based desolvation." Journal of Computational Chemistry **28**(6): 1145-1152.

Jantsch, A., L. Nieto, J. Gertsch, J. Rodríguez-Salarichs, R. Matesanz, J. Jiménez-Barbero, J. F. Díaz, Á. Canales and K. H. Altmann (2016). "Synthesis, biological profiling and determination of the tubulin-bound conformation of 12-aza-epothilones (azathilones)." Molecules **21**(8).

Jayalakshmi, V., T. Biet, T. Peters and N. R. Krishna (2004). "Refinement of the conformation of UDP-galactose bound to galactosyltransferase using the STD NMR intensity-restrained CORCEMA optimization." J Am Chem Soc **126**: 8610-8611.

Jayalakshmi, V. and N. R. Krishna (2002). "Complete relaxation and conformational exchange matrix (CORCEMA) analysis of intermolecular saturation transfer effects in reversibly forming ligand-receptor complexes." J Magn Reson **155**: 106-118.

Jayalakshmi, V. and N. Rama Krishna (2004). "CORCEMA refinement of the bound ligand conformation within the protein binding pocket in reversibly forming weak complexes using STD-NMR intensities." Journal of Magnetic Resonance **168**(1): 36-45.

Jeener, J., B. H. Meier, P. Bachmann and R. R. Ernst (1979). "Investigation of exchange processes by two-dimensional NMR spectroscopy." The Journal of Chemical Physics **71**(11): 4546-4553.

Jeener, J., B. H. Meier, P. Bachmann and R. R. Ernst (1979). "A two-dimensional nuclear overhauser experiment with pure phase absorption in four quadrants." J. Chem. Phys **71**: 4546-4553.

Jenkins, C., R. Samudrala, I. Anderson, B. P. Hedlund, G. Petroni, N. Michailova, N. Pinel, R. Overbeek, G. Rosati and J. T. Staley (2002). "Genes for the cytoskeletal protein tubulin in the bacterial genus *Prostheco bacter*." Proc Natl Acad Sci U S A **99**: 17049-17054.

Jimenez-Barbero, J., J. L. Asensio, F. J. Canada and A. Poveda (1999). "Reviews-Free and protein-bound carbohydrate structures." Current Opinion in Structural Biology **9**(5): 549-555.

Jiménez-Barbero, J., A. Canales, P. T. Northcote, R. M. Buey, J. M. Andreu and J. F. Díaz (2006). "NMR determination of the bioactive conformation of peloruside a bound to microtubules." J Am Chem Soc **128**: 8757-8765.

Jordan, M. A. (2002). "Mechanism of action of antitumor drugs that interact with microtubules and tubulin." Curr Med Chem Anti-Canc Agents **2**: 1-17.

Jordan, M. A., K. Wendell, S. Gardiner, W. B. Derry, H. Copp and L. Wilson (1996). "Mitotic block induced in HeLa cells by low concentrations of paclitaxel (Taxol) results in abnormal mitotic exit and apoptotic cell death." Cancer Res **56**: 816-825.

Jordan, M. A. and L. Wilson (2004). "Microtubules as a target for anticancer drugs." Nat Rev Cancer **4**: 253-265.

Jorgensen, W. L., D. S. Maxwell and J. Tirado-Rives (1996). "Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids." Journal of the American Chemical Society **118**(45): 11225-11236.

Jorgensen, W. L. and J. Tiradorives (1988). "The OPLS potential functions for proteins - energy minimizations for crystals of cyclic-peptides and crambin." JACS **110**: 1657-1666.

Kaminski, G. A., R. A. Friesner, J. Tirado-Rives and W. L. Jorgensen (2001). "Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides." J. Phys. Chem. B **105**: 6474-6487.

Kavallaris, M. (2010). "Microtubules and resistance to tubulin-binding agents." Nat Rev Cancer **10**(3): 194-204.

Keller, R. M., D. Picot and K. Wüthrich (1979). "Individual assignments of the heme resonances in the 360 MHz ¹H NMR spectra of cytochrome c-557 from

Crithidia oncopelti." Biochimica et Biophysica Acta (BBA) - Protein Structure **580**(2): 259-265.

Keserû, G. M. and I. Kolossváry (2001). "Fully flexible low-mode docking: application to induced fit in HIV integrase." Journal of the American Chemical Society **123**(50): 12708-12709.

Kirkpatrick, S., C. D. Gelatt and M. P. Vecchi (1983). "Optimization by Simulated Annealing." Science **220**(4598): 671.

Kolossváry, I. and W. C. Guida (1996). "Low mode search. An efficient, automated computational method for conformational analysis: Application to cyclic and acyclic alkanes and cyclic peptides." Journal of the American Chemical Society **118**(21): 5011-5019.

Krishna, N. and V. Jayalakshmi (2008). Quantitative Analysis of STD-NMR Spectra of Reversibly Forming Ligand–Receptor Complexes. Bioactive Conformation II. T. Peters, Springer Berlin / Heidelberg. **273**: 15-54.

Krishna, N. R., D. G. Agresti, J. D. Glickson and R. Walter (1978). "Solution conformation of peptides by the intramolecular nuclear Overhauser effect experiment. Study of valinomycin-K⁺." Biophysical Journal **24**(3): 791-814.

Krishna, N. R. and L. Berliner (1999). Structure computation and dynamics in protein NMR, Springer Science & Business Media.

Krishna, N. R., G. Goldstein and J. D. Glickson (1980). "A general multistate model for the analysis of hydrogen-exchange kinetics." Biopolymers **19**(11): 2003-2020.

Krishna, N. R. and H. N. B. Moseley (2002). "Complete Relaxation and Conformational Exchange Matrix (CORCEMA) Analysis of NOESY Spectra of Reversibly Forming Ligand-Receptor Complexes Application to Transferred NOESY." Biological Magnetic Resonance: Structure Computation and Dynamics in Protein NMR: 223-307.

Kumar, A., R. R. Ernst and K. Wüthrich (1980). "A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules." Biochemical and biophysical research communications **95**(1): 1-6.

Kuntz, D. A., A. Ghavami, B. D. Johnston, B. M. Pinto and D. R. Rose (2005). "Crystallographic analysis of the interactions of *Drosophila melanogaster* Golgi α -mannosidase II with the naturally occurring glycomimetic salacinol and its analogues." Tetrahedron: Asymmetry **16**(1): 25-32.

Landau, L. D. and E. M. Lifshitz (2013). Statistical Physics, Elsevier Science.

Lee, J. C., R. P. Frigon and S. N. Timasheff (1973). "The chemical characterization of calf brain microtubule protein subunits." J Biol Chem **248**: 7253-7262.

Lee, W. and N. R. Krishna (1992). "Influence of conformational exchange on the 2D NOESY spectra of biomolecules existing in multiple conformations." Journal of Magnetic Resonance (1969) **98**(1): 36-48.

Lepre, C. A., J. M. Moore and J. W. Peng (2004). "Theory and Applications of NMR-Based Screening in Pharmaceutical Research." Chemical Reviews **104**(8): 3641-3676.

Lomax, R. G. and D. L. Hahs-Vaughn (2013). Statistical Concepts: A Second Course, Taylor & Francis.

Lowe, J., H. Li, K. H. Downing and E. Nogales (2001). "Refined structure of alpha beta-tubulin at 3.5 Å resolution." J Mol Biol **313**: 1045-1057.

Maggioni, A., M. von Itzstein, J. Tiralongo and T. Haselhorst (2008). "Detection of Ligand Binding to Nucleotide Sugar Transporters by STD NMR Spectroscopy." ChemBioChem **9**(17): 2784-2786.

Mani, S., M. Macapinlac, S. Goel, D. Verdier-Pinard, T. Fojo, M. Rothenberg and D. Colevas (2004). "The clinical development of new mitotic inhibitors that stabilize the microtubule." Anticancer Drugs **15**: 553-558.

Marcelo, F., S. Huecas, L. B. Ruiz-Ávila, F. J. Cañada, A. Perona, A. Poveda, S. Martín-Santamaría, A. Morreale, J. Jiménez-Barbero and J. M. Andreu (2013). "Interactions of Bacterial Cell Division Protein FtsZ with C8-Substituted Guanine Nucleotide Inhibitors. A Combined NMR, Biochemical and Molecular Modeling Perspective." Journal of the American Chemical Society **135**(44): 16418-16428.

Marchetti, R., S. Perez, A. Arda, A. Imberty, J. Jimenez-Barbero, A. Silipo and A. Molinaro (2016). "'Rules of Engagement' of Protein–Glycoconjugate Interactions: A Molecular View Achievable by using NMR Spectroscopy and Molecular Modeling." ChemistryOpen **5**(4): 274-296.

Markley, J. L. and S. J. Opella (1997). Biological NMR Spectroscopy, Oxford University Press.

Matesanz, R., I. Barasoain, C. Yang, L. Wang, X. Li, C. De Ines, C. Coderch, F. Gago, J. Jiménez-Barbero, J. M. Andreu, W. Fang and J. F. Díaz (2008). "Optimization of taxane binding to microtubules. Binding affinity decomposition and incremental construction of a high-affinity analogue of paclitaxel." Chem Biol **15**: 573-585.

Matesanz, R., J. Rodríguez-Salarichs, B. Pera, Á. Canales, J. M. Andreu, J. Jiménez-Barbero, W. Bras, A. Nogales, W.-S. Fang and J. F. Díaz (2011). "Modulation of microtubule interprotofilament interactions by modified taxanes." Biophysical journal **101**(12): 2970-2980.

Matesanz, R., C. Trigili, J. Rodríguez-Salarichs, I. Zanardi, B. Pera, A. Nogales, W. S. Fang, J. Jímenez-Barbero, Á. Canales, I. Barasoain, I. Ojima and J. F. Díaz (2015). "Taxanes with high potency inducing tubulin assembly overcome

tumoural cell resistances." Bioorganic and Medicinal Chemistry **22**(18): 5078-5090.

Maudsley, A. A., L. Moller and R. R. Ernst (1977). "Cross-correlation of spin-decoupled NMR spectra by heteronuclear two-dimensional spectroscopy." Journal of Magnetic Resonance (1969) **28**(3): 463-469.

Mayer, M. and T. L. James (2004). "NMR-Based Characterization of Phenothiazines as a RNA Binding Scaffold." Journal of the American Chemical Society **126**(13): 4453-4460.

Mayer, M. and B. Meyer (1999). "Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy." Angewandte Chemie International Edition **38**(12): 1784-1788.

McCoy, A. J., R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read (2007). "Phaser crystallographic software." Journal of Applied Crystallography **40**(Pt 4): 658-674.

Meagher, K. L., L. T. Redman and H. A. Carlson (2003). "Development of polyphosphate parameters for use with the AMBER force field." J. Comput. Chem. **24**: 1016-1025.

Meier, B. H. and R. R. Ernst (1979). "Elucidation of chemical exchange networks by two-dimensional NMR spectroscopy: the heptamethylbenzenonium ion." Journal of the American Chemical Society **101**(21): 6441-6442.

Meyer, B. and T. Peters (2003). "NMR Spectroscopy techniques for screening and identifying ligand binding to protein receptors." Angewandte Chemie-International Edition **42**: 864-890.

Michalewicz, Z. (2013). Genetic algorithms+ data structures= evolution programs, Springer Science & Business Media.

Mohamadi, F., N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrickson and W. C. Still (1990). "MacroModel an integrated software system for modeling organic and bioorganic molecules using molecular mechanics." J. Comput. Chem. **11**: 440-467.

Morales, J. L. and J. N. Y (2000). "Automatic preconditioning by limited memory Quasi-Newton updating." SIAM J. Optim **10**: 1079-1096.

Morris, G. M., D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew and A. J. Olson (1998). "Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function." Journal of Computational Chemistry **19**(14): 1639-1662.

Moseley, H. N. B., E. V. Curto and N. R. Krishna (1995). "Complete relaxation and conformational exchange matrix (CORCEMA) analysis of NOESY spectra of interacting systems - 2-dimensional transferred NOESY." J Mag Reson Ser B **108**: 243-261.

Moseley, H. N. B., W. Lee, C. H. Arrowsmith and N. R. Krishna (1997). "Quantitative determination of conformational, dynamic, and kinetic parameters of a ligand-protein/DNA complex from a complete relaxation and conformational exchange matrix analysis of intermolecular transferred NOESY." Biochemistry **36**(18): 5293-5299.

Muñoz-García, J. C., E. Chabrol, R. R. Vivès, A. Thomas, J. L. de Paz, J. Rojo, A. Imberty, F. Fieschi, P. M. Nieto and J. Angulo (2015). "Langerin–Heparin Interaction: Two Binding Sites for Small and Large Ligands As Revealed by a Combination of NMR Spectroscopy and Cross-Linking Mapping Experiments." Journal of the American Chemical Society **137**(12): 4100-4110.

Nagayama, K., K. Wüthrich, P. Bachmann and R. R. Ernst (1977). "Two-dimensional NMR Spectroscopy." Naturwissenschaften **64**(11): 581-583.

Neuhaus, B. and M. P. Williamson (1989). The Nuclear Overhauser Effect in Structural and Conformational Analysis. New York, VCH.

Nietlispach, D. and A. Gautier (2011). "Solution NMR studies of polytopic α -helical membrane proteins." Engineering and design / Membranes **21**(4): 497-508.

O'Boyle, N. M., M. Banck, C. A. James, C. Morley, T. Vandermeersch and G. R. Hutchison (2011). "Open Babel: An open chemical toolbox." Journal of cheminformatics **3**(1): 1-14.

Overhauser, A. W. (1953). "Polarization of nuclei in metals." Physical Review **92**(2): 411.

Pera, B., I. Barasoain, A. Pantazopoulou, A. Canales, R. Matesanz, J. Rodriguez-Salarichs, L. F. García-Fernandez, V. Moneo, J. Jiménez-Barbero and C. M. Galmarini (2013). "New Interfacial Microtubule Inhibitors of Marine Origin, PM050489/PM060184, with Potent Antitumor Activity and a Distinct Mechanism." ACS chemical biology **8**(9): 2084-2094.

Pera, B., I. Barasoain, A. Pantazopoulou, A. Canales, R. Matesanz, J. Rodriguez-Salarichs, L. F. García-Fernandez, V. Moneo, J. Jiménez-Barbero, C. M. Galmarini and others (2013). "New Interfacial Microtubule Inhibitors of Marine Origin, PM050489/PM060184, with Potent Antitumor Activity and a Distinct Mechanism." ACS chemical biology **8**(9): 2084-2094.

Peters, T. and B. Meyer (2001). Method for detecting biologically active compounds from compound libraries, Google Patents.

Poveda, A., J. L. Asensio, M. MartínPastor and J. JiménezBarbero (1997). "Solution conformation and dynamics of a tetrasaccharide related to the Lewis(x) antigen deduced by NMR relaxation measurements." Journal of Biomolecular Nmr **10**: 29-43.

Poveda, A. and J. Jimenez-Barbero (1998). "NMR studies of carbohydrate-protein interactions in solution." Chemical Society Reviews **27**(2): 133-144.

Prota, A. E., K. Bargsten, J. F. Diaz, M. Marsh, C. Cuevas, M. Liniger, C. Neuhaus, J. M. Andreu, K.-H. Altmann and M. O. Steinmetz (2014). "A new tubulin-binding site and pharmacophore for microtubule-destabilizing anticancer drugs." Proceedings of the National Academy of Sciences **111**(38): 13817-13821.

Prota, A. E., K. Bargsten, D. Zurwerra, J. J. Field, J. F. Díaz, K.-H. Altmann and M. O. Steinmetz (2013). "Molecular Mechanism of Action of Microtubule-Stabilizing Anticancer Agents." Science **339**(6119): 587-590.

Prota, A. E., M. M. Magiera, M. Kuijpers, K. Bargsten, D. Frey, M. Wieser, R. Jaussi, C. C. Hoogenraad, R. A. Kammerer, C. Janke and M. O. Steinmetz (2013). "Structural basis of tubulin tyrosination by tubulin tyrosine ligase." The Journal of Cell Biology **200**(3): 259.

Ravelli, R. B., B. Gigant, P. A. Curmi, I. Jourdain, S. Lachkar, A. Sobel and M. Knossow (2004). "Insight into tubulin regulation from a complex with colchicine and a stathmin-like domain." Nature **428**: 198-202.

Richard, E. (1969). Impulse resonance spectrometer including a time averaging computer and fourier analyzer, Google Patents.

Rotkiewicz, P. and J. Skolnick (2008). "Fast procedure for reconstruction of full-atom protein models from reduced representations." Journal of computational chemistry **29**(9): 1460-1465.

Sánchez, S. A., J. E. Brunet, D. M. Jameson, R. Lagos and O. Monasterio (2004). "Tubulin equilibrium unfolding followed by time-resolved fluorescence and fluorescence correlation spectroscopy." Protein Science : A Publication of the Protein Society **13**(1): 81-88.

Schlegel, H. B. (1982). "Optimization of equilibrium geometries and transition structures." Journal of Computational Chemistry **3**(2): 214-218.

Schlegel, H. B. (2011). "Geometry optimization." Wiley Interdisciplinary Reviews: Computational Molecular Science **1**(5): 790-809.

Schlieper, D., M. A. Oliva, J. M. Andreu and J. Lowe (2005). "Structure of bacterial tubulin BtubA/B: Evidence for horizontal gene transfer." Proc Natl Acad Sci U S A.

Shabbits, J. A., R. Krishna and L. D. Mayer (2001). "Molecular and pharmacological strategies to overcome multidrug resistance." Expert Rev Anticancer Ther **1**: 585-594.

Shah, N., D. A. Kuntz and D. R. Rose (2003). "Comparison of Kifunensine and 1-Deoxymannojirimycin Binding to Class I and II α -Mannosidases Demonstrates Different Saccharide Distortions in Inverting and Retaining Catalytic Mechanisms." Biochemistry **42**(47): 13812-13816.

Shapiro, A. and M. Fan (1995). "On Eigenvalue Optimization." SIAM Journal on Optimization **5**(3): 552-569.

Siegel, R. L., K. D. Miller and A. Jemal (2015). "Cancer statistics, 2015." CA: A Cancer Journal for Clinicians **65**(1): 5-29.

Smith, G. M. (1979). "Proton nuclear magnetic resonance studies of *Rhodospirillum rubrum* cytochrome c2." Biochemistry **18**(8): 1628-1634.

Solomon, I. (1955). "Relaxation processes in a system of two spins." Physical Review **99**(2): 559.

Still, W. C., A. Tempczyk, R. C. Hawley and T. Hendrickson (1990). "Semianalytical treatment of solvation for molecular mechanics and dynamics." J Am Chem Soc **112**: 6127-6129.

Stock, D., C. Gibbons, I. Arechaga, A. G. W. Leslie and J. E. Walker (2000). "The rotary mechanism of ATP synthase." Current opinion in structural biology **10**(6): 672-679.

Stock, D., A. G. W. Leslie and J. E. Walker (1999). "Molecular architecture of the rotary motor in ATP synthase." Science **286**(5445): 1700-1705.

Strehlow, H. and W. Knoche (1977). Fundamentals of Chemical Relaxation. Weinheim, Germany, Verlag Chemie: 101-105.

Szczepina, B. M. a. M. G. a. B. D. W. b. P. (2011). "Investigation of the binding of a carbohydrate-mimetic peptide to its complementary anticarbohydrate antibody by STD-NMR spectroscopy and molecular-dynamics simulations." Chemistry - A European Journal **17**(41): 11446-11455.

Szczepina, M. G., D. W. Bleile and B. M. Pinto (2011). "Investigation of the Binding of a Carbohydrate-Mimetic Peptide to its Complementary Anticarbohydrate Antibody by STD-NMR Spectroscopy and Molecular-Dynamics Simulations." Chemistry – A European Journal **17**(41): 11446-11455.

Tanaka, J. and T. Higa (1996). "Zampanolide, a new cytotoxic macrolide from a marine sponge." Tetrahedron Letters **37**: 5535-5538.

Tang, Y., J. Rodríguez-Salarichs, Y. Zhao, P. Cai, J. Estévez-Gallego, F. Balaguer-Pérez, M. Redondo Horcajo, D. Lucena-Agell, I. Barasoain, J. F. Díaz and W.-S. Fang (2017). "Modification of C-seco taxoids through ring tethering and substituent replacement leading to effective agents against tumor drug resistance mediated by β III-Tubulin and P-glycoprotein (P-gp) overexpressions." European Journal of Medicinal Chemistry.

Thomas, E., J. Tabernero, M. Fornier, P. Conte, P. Fumoleau, A. Lluch, L. T. Vahdat, C. A. Bunnell, H. A. Burris, P. Viens, J. Baselga, E. Rivera, V. Guarneri, V. Poulart, J. Klimovsky, D. Lebwohl and M. Martin (2007). "Phase II clinical trial

of ixabepilone (BMS-247550), an epothilone B analog, in patients with taxane-resistant metastatic breast cancer." J Clin Oncol **25**: 3399-3406.

Torre, L. A., F. Bray, R. L. Siegel, J. Ferlay, J. Lortet-Tieulent and A. Jemal (2015). "Global cancer statistics, 2012." CA: A Cancer Journal for Clinicians **65**(2): 87-108.

Trott, O. and A. J. Olson (2010). "AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading." Journal of Computational Chemistry **31**(2): 455-461.

Vanommeslaeghe, K., E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov and A. D. Mackerell (2010). "CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields." Journal of Computational Chemistry **31**(4): 671-690.

Venkitakrishnan, R. P., O. Benard, M. Max and J. L. Markley (2012). "Use of NMR Saturation Transfer Difference Spectroscopy to Study Ligand Binding to Membrane Proteins." Methods in molecular biology (Clifton, N.J.) **914**: 47-63.

Viegas, A., J. Manso, F. L. Nobrega and E. J. Cabrita (2011). "Saturation-Transfer Difference (STD) NMR: A Simple and Fast Method for Ligand Screening and Characterization of Protein Binding." Journal of Chemical Education **88**(7): 990-994.

Vogtherr, M. and T. Peters (2000). "Application of NMR Based Binding Assays to Identify Key Hydroxy Groups for Intermolecular Recognition." Journal of the American Chemical Society **122**(25): 6093-6099.

Wagstaff, J. L., S. L. Taylor and M. J. Howard (2013). "Recent developments and applications of saturation transfer difference nuclear magnetic resonance (STD NMR) spectroscopy." Molecular BioSystems **9**(4): 571-577.

Wang, J., R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case (2004). "Development and testing of a general amber force field." Journal of Computational Chemistry **25**(9): 1157-1174.

Warshel, A. and M. Levitt (1976). "Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme." Journal of Molecular Biology **103**(2): 227-249.

Weisenberg, R. C., G. G. Borisy and E. W. Taylor (1968). "The colchicine-binding protein of mammalian brain and its relation to microtubules." Biochemistry **7**: 4466-4479.

Wen, B. M. a. X. a. Y. Y. a. K. D. A. b. R. D. R. b. P. (2005). "A combined STD-NMR/molecular modeling protocol for predicting the binding modes of the glycosidase inhibitors kifunensine and salacinol to golgi α -mannosidase II." Biochemistry **44**(18): 6729-6737.

Willcott, M. R., R. E. Lenkinski and R. E. Davis (1972). "Interpretation of the pseudocontact model for nuclear magnetic resonance shift reagents. I. Agreement factor, R." Journal of the American Chemical Society **94**(5): 1742-1744.

Wright, S. and J. Nocedal (1999). "Numerical optimization." Springer Science **35**: 67-68.

Wüthrich, K., G. Wider, G. Wagner and W. Braun (1982). "Sequential resonance assignments as a basis for determination of spatial protein structures by high resolution proton nuclear magnetic resonance." Journal of molecular biology **155**(3): 311-319.

Yan, J., A. D. Kline, H. Mo, M. J. Shapiro and E. R. Zartler (2003). "The effect of relaxation on the epitope mapping by saturation transfer difference NMR." Journal of Magnetic Resonance **163**(2): 270-276.

Yanamala, N., A. Dutta, B. Beck, B. Van Fleet, K. Hay, A. Yazbak, R. Ishima, A. Doemling and J. Klein-Seetharaman (2010). "NMR-Based Screening of Membrane Protein Ligands." Chemical Biology & Drug Design **75**(3): 237-256.

Yang, C., I. Barasoain, X. Li, R. Matesanz, R. Liu, F. J. Sharom, J. F. Díaz and W. Fang (2007). "Overcoming Tumor Drug Resistance Mediated by P-glycoprotein Overexpression with high affinity taxanes: A SAR study of C-2 Modified 7-Acyl-10-Deacetyl Cephalomannines." Chem Med Chem **2**: 691-701.

Yuan, B. M. a. Y. a. W. X. a. S. D. A. R. b. P. (2005). "Exploring the mechanism of binding of UDP-galactopyranose to UDP-galactopyranose mutase by STD-NMR spectroscopy and molecular modeling." Biochemistry **44**(43): 14080-14089.